stochastic
processes
and their

www.elsevier.com/locate/spa

The loop erased exit path and the metastability of a biased vote process

Olivier Catoni[a,*], Dayue Chen[b], Jun Xie[b]

[a] *CNRS, UMR 7599, Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 6, Tour 56-66, 3ième étage, 4 Place Jussieu, F-75 252 Paris, Cedex 05, France*

[b] *Department of Probability and Statistics, Peking University, Beijing 100871, People's Republic of China*

Received 7 January 1999; received in revised form 27 September 1999

Abstract

The reduction method provides an algorithm to compute large deviation estimates of (possibly nonreversible) Markov processes with exponential transition rates. It replaces the original graph minimisation equations of Freidlin and Wentzell by more tractable path minimisation problems. When applied to study the metastability of the dynamics, it gives a large deviation principle for the loop erased exit path from the metastable state. To illustrate this, we study a biased majority vote process generalising the one introduced in Chen (1997. J. Statist. Phys. 86 (3/4), 779–802). We show that this nonreversible dynamics has a two well potential with a unique metastable state, we give an upper bound for its relaxation time, and show that for small enough values of the bias the exit path is typically different at low temperature from the typical exit paths of the Ising model.

MSC: 60F10; 60Jxx; 82C05

Keywords: Finite Markov chains with exponential transitions; Metastability; Biased majority vote process; Large deviations

1. The model

We consider on some finite state space E a family of infinitesimal generators L_β indexed by a positive real "inverse temperature" parameter β.

$$(L_\beta f)(\sigma) = \sum_{\sigma' \neq \sigma} c_\beta(\sigma, \sigma')(f(\sigma') - f(\sigma)),$$

where

$$\lim_{\beta \to +\infty} -\frac{1}{\beta} \log c_\beta(\sigma, \sigma') = V(\sigma, \sigma') \in \mathbb{R}_+ \cup \{+\infty\}. \quad (1)$$

* Corresponding author.

E-mail addresses: catoni@ccr.jussieu.fr (O. Catoni), dayue@sxx0.math.pku.edu.cn (D. Chen)

0304-4149/00/$ - see front matter
PII: S0304-4149(99)00096-4

We let P_t^β be the semi-group

$$\sum_{n=0}^{+\infty} \frac{t^n}{n!} L_\beta^n$$

generated by L_β and consider on the space $D([0,+\infty[,E)$ of right continuous trajectories with left limits the canonical process $(Z_t)_{t\in\mathbb{R}_+}$ and the family of probability measures $\bar{\mathbb{P}}_\sigma^\beta \in \mathscr{M}_+^1(D([0,+\infty[,E))$ of Markov processes starting from σ with semi-groups P_t^β.

We know from the matrix tree theorem (Kirchhoff, 1847; Freidlin and Wentzell, 1984; Chen et al., 1996) that the invariant probability distribution μ_β of P_t^β can be expressed with the help of some special families of graphs. More precisely, using for any graph $g \subset E\times E$ the functional notation $g(\sigma)=\{\sigma' \in E \mid (\sigma,\sigma') \in g\}$, we can define for any subset $A \subsetneq E$

$$G(A)=\{g\subset E\times E; |g(\sigma)|=\mathbf{1}(\sigma\notin A) \text{ and } g(C)\neq C \text{ for all } C\subset E\setminus A\}.$$

Then considering the weights

$$c_\beta(g)=\prod_{(\sigma,\sigma')\in g} c_\beta(\sigma,\sigma'),$$

we have

$$\mu_\beta(\sigma)=\frac{1}{Z(\beta)}\sum_{g\in G(\{\sigma\})} c_\beta(g),$$

and therefore we can define the virtual energy of $(\mu_\beta)_{\beta\in\mathbb{R}_+}$ to be

$$U(\sigma)=\lim_{\beta\to+\infty} -\frac{1}{\beta}\log \mu_\beta(\sigma).$$

More generally, we will put for any $A\subset E$

$$U(A)=\lim_{\beta\to+\infty} -\frac{1}{\beta}\log \mu_\beta(A)=\min\{U(\sigma);\sigma\in A\},$$

and let $\bar{\tau}(A)$ be the exit time from A:

$$\bar{\tau}(A)=\inf\{t\geqslant 0;\; Z_t\notin A\}.$$

The study of the metastability phenomenon at low temperature is concerned with the following question: Pick up some ground state $g\in \arg\min U$ and ask what are the most remote states g' from g, in the sense that

$$\lim_{\beta\to+\infty} \frac{1}{\beta}\log \mathbb{E}_{g'}^\beta(\bar{\tau}(E\setminus\{g\}))$$

is the largest. For such a g' (necessarily different from g), describe the distribution of the exit path

$$\mathscr{L}_{g'}^\beta((Z_{\bar{\theta}+t})_{0\leqslant t\leqslant \bar{\tau}-\bar{\theta}}),$$

where

$$\bar{\tau}=\inf\{t\mid Z_t=g\},$$
$$\bar{\theta}=\sup\{t\leqslant\bar{\tau}; Z_t=g'\},$$

when β tends to $+\infty$. More precisely, after we have made the link between continuous and discrete time dynamics, we will be able to apply the results of Catoni and Cerf (1997) and to state a large deviation principle for the order in which the states are visited by the exit path: to explain what we have in mind, let us define the random times

$$\bar{v}_0 = \bar{\theta},$$
$$\bar{v}_k = \inf\{\bar{\tau} \geqslant t > \bar{v}_{k-1}; Z_t \notin \{Z_s, \bar{\theta} \leqslant s < t\}\},$$

with the convention that $\inf \emptyset = +\infty$ and $Z_{+\infty} = \Delta \notin E$ where Δ is some "coffin state". These are the succession of times when the support of the exit path increases. We will state a large deviation estimate for the sequence of newly visited states, that is for

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \bar{\mathbb{P}}^{\beta}_{g'}((Z_{\bar{v}_k})_{k=0}^{+\infty} \in \mathscr{A}),$$

for any measurable subset $\mathscr{A} \subset (E \cup \{\Delta\})^{\mathbb{N}}$. Moreover for any $\sigma \in E$ we will compute the following large deviation rate function of the local time of the exit path:

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \bar{\mathbb{E}}^{\beta}_{g'} \left(\int_{\bar{\theta}}^{\bar{\tau}} \mathbf{1}(Z_s = \sigma)\, \mathrm{d}s \,|\, \sigma \in \{Z_s; \bar{\theta} \leqslant s \leqslant \bar{\tau}\} \right), \quad \sigma \in E \setminus \{g, g'\}.$$

We will also describe a new approach to the exit path through a new "reduction algorithm". This will lead us to define a loop eraser map $\Gamma((Z_{\bar{\theta}+t})_{t=0}^{\bar{\tau}-\bar{\theta}})$ that will erase most of the loops in the exit path and will produce a pruned discrete time path valued in some finite set of possible "reduction paths". We will establish a large deviation principle for the loop erased exit path, that is for

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{P}^{\beta}_{g'}(\Gamma((Z_{\bar{\theta}+t})_{t=0}^{\bar{\tau}-\bar{\theta}}) = \gamma).$$

The reduction algorithm will give us an efficient way to compute explicitly the rate function in many situations and hopefully to find the reduction paths of minimal cost. In the last section, we apply these ideas to the study of a nonreversible dynamics, the biased majority vote process on the torus.

2. Discrete dynamics

A first technical step towards an "explicit" solution of the previous questions based on probabilistic tools is to introduce a discrete time dynamics.

It is easy here because the state space E is finite and therefore the rates $c_\beta(\sigma, \sigma')$ are bounded for any fixed value of β. Let us consider for each β some constant K_β such that

$$\sup_{\sigma} \sum_{\sigma' \neq \sigma} c_\beta(\sigma, \sigma') \leqslant \frac{K_\beta}{2}.$$

(The factor $1/2$ on the right-hand side will be useful to get Eq. (2) below.)

As we have assumed that function V of Eq. (1) was nonnegative, we can and will choose $\{K_\beta; \beta \in \mathbb{R}_+\}$ such that

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log(K_\beta) = 0.$$

If we let $p_\beta = (I + (1/K_\beta)L_\beta)$, we obtain a Markov matrix, and it is legitimate to consider the canonical process $(X_n)_{n\in\mathbb{N}}$ on $E^{\mathbb{N}}$ equipped with the distribution $\mathbb{P}_\sigma^\beta$ of the time homogeneous Markov chain starting from state σ with transition matrix p_β.

On the space of right continuous trajectories with left limits on the integers $D([0,\infty[,\mathbb{N})$, let $\mathbb{Q}$ be the distribution of the Poisson process with intensity $K_\beta t$ starting from 0, and let $(N_t)_{t\in\mathbb{R}_+}$ be the canonical process on this space. We can extend the canonical processes $(X_n)_{n\in\mathbb{N}}$ and $(N_t)_{t\in\mathbb{R}_+}$ in an obvious way to the product space $\Omega = E^{\mathbb{N}} \times D([0,+\infty[,\mathbb{N})$, putting for any $\omega = (\omega_1,\omega_2) \in \Omega$ $X_n(\omega) \overset{\text{def}}{=} X_n(\omega_1)$ and $N_t(\omega) \overset{\text{def}}{=} N_t(\omega_2)$.

Then on the product space $\Omega = E^{\mathbb{N}} \times D([0,+\infty[,\mathbb{N})$, we can consider the process $(X_{N_t})_{t\in\mathbb{R}_+}$ whose distribution under the product measure $\mathbb{P}_\sigma^\beta \otimes \mathbb{Q}$ is equal to $\bar{\mathbb{P}}_\sigma^\beta$. In the following, we will work with the realisation $Z_t = X_{N_t}$ and use $\mathbb{E}_\sigma^\beta$ to denote the expectation with respect to $\mathbb{P}_\sigma^\beta \otimes \mathbb{Q}$.

We claim that the features of the exit path that we want to analyse can be computed on the discrete time dynamics $(X_n)_{n\in\mathbb{N}}$. Indeed we have the following straightforward links between $((X_n)_{n\in\mathbb{N}}, \Omega, \mathbb{P}_\sigma^\beta \otimes \mathbb{Q})$ and $((Z_t)_{t\in\mathbb{R}_+}, \Omega, \mathbb{P}_\sigma^\beta \otimes \mathbb{Q})$ (note that we work with the realisation $Z_t = X_{N_t}$ and define X_n and Z_t on the same product probability space Ω):

- The large deviation rates of the transitions are the same: for any $\sigma \neq \sigma' \in E$

$$\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{P}_\sigma^\beta(X_1 = \sigma') = -V(\sigma,\sigma').$$

  Moreover, as we introduced a factor 1/2 in the time renormalisation, we are sure that for any $\sigma \in E$

$$\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{P}_\sigma^\beta(X_1 = \sigma) = 0, \tag{2}$$

  so that we will extend V to the diagonal in the following and put for any $\sigma \in E$

$$V(\sigma,\sigma) = 0.$$

- The invariant probability measure μ_β of $\bar{\mathbb{P}}_\sigma^\beta$ is also the invariant probability measure of $\mathbb{P}_\sigma^\beta$.
- Exit times and local times satisfy the following relations: if for any $A \subsetneq E$

$$\tau(A) = \inf\{n \geqslant 0; X_n \notin A\},$$
$$\bar{\tau}(A) = \inf\{t \geqslant 0; Z_t \notin A\}$$

  then

$$\mathbb{E}_\sigma^\beta(\bar{\tau}(A) \mid (X_n)_{n\in\mathbb{N}}) = K_\beta \tau(A),$$

$$\mathbb{E}_\sigma^\beta\left(\int_0^{\bar{\tau}(A)} \mathbf{1}(Z_s = \sigma')\mathrm{d}s \mid (X_n)_{n\in\mathbb{N}}\right) = K_\beta \sum_{n=0}^{\tau(A)} \mathbf{1}(X_n = \sigma'), \quad \sigma,\sigma' \in A,$$

  and similarly for any $\sigma \in E \setminus \{g,g'\}$,

$$\mathbb{E}_{g'}^\beta\left(\int_{\bar{\theta}}^{\bar{\tau}} \mathbf{1}(Z_s = \sigma)\mathrm{d}s \mid \sigma \in \{Z_s; \bar{\theta} \leqslant s \leqslant \bar{\tau}\}\right)$$

$$= \mathbb{E}^{\beta}_{g'} \left(\int_{\bar\theta}^{\bar\tau} \mathbf{1}(Z_s = \sigma) \, \mathrm{d}s \mid \sigma \in \{X_n; \theta \leqslant n \leqslant \tau\} \right)$$

$$= K_\beta \, \mathbb{E}^{\beta}_{g'} \left(\sum_{n=\theta}^{\tau} \mathbf{1}(X_n = \sigma) \mid \sigma \in \{X_n; \theta \leqslant n \leqslant \tau\} \right),$$

where

$$\tau = \inf\{n \in \mathbb{N}; X_n = g\},$$
$$\theta = \sup\{n \leqslant \tau; X_n = g'\}.$$

- A small (unessential) difference is that $Z_{\bar\theta} = X_{\theta+1}$, because we are working with a right continuous realization of the Poisson process (N_t). Therefore the relation between the supports of the continuous time and the discrete time exit paths is

$$\{Z_s; \bar\theta \leqslant s \leqslant \bar\tau\} = \{X_n; \theta < n \leqslant \tau\}.$$

One nice feature of the discrete time dynamics is the possibility to introduce the modified stopping time

$$\tau'(A) = \inf\{n > 0 \mid X_n \notin A\}.$$

Note that $\tau'(A)$ is not measurable with respect to $(Z_t) = (X_{N_t})$, because (N_t) is not measurable with respect to (Z_t): indeed, when $X_{n+1} = X_n$, the jump time of (N_t) from value n to value $n+1$ cannot be seen on (Z_t). Therefore there is no corresponding notion of modified exit time in continuous time.

The distribution of the exit path of $(X_n)_{n\in\mathbb{N}}$ can easily be proved to be also the distribution of a simpler object involving only stopping times:

Proposition 2.1. *The following equality between distributions holds*:

$$\mathscr{L}^{\beta}_{g'}((X_{\theta+n})_{0\leqslant n\leqslant \tau-\theta}) = \mathscr{L}^{\beta}_{g'}((X_n)_{0\leqslant n\leqslant \tau'(E\setminus\{g,g'\})} \mid X_{\tau'(E\setminus\{g,g'\})} = g).$$

Proof. For any sequence $(x_0,\dots,x_r)$ such that $x_0=g'$, $x_r=g$ and $x_1,\dots,x_{r-1} \in E\setminus\{g,g'\}$, we have

$$\mathbb{P}^{\beta}_{g'}((X_{\theta+n})_{0\leqslant n\leqslant\tau-\theta} = (x_0,\dots,x_r))$$
$$= \sum_{k=0}^{+\infty} \mathbb{P}^{\beta}_{g'}(\theta = k, (X_{\theta+n})_{0\leqslant n\leqslant\tau-\theta} = (x_0,\dots x_r))$$
$$= \sum_{k=0}^{+\infty} \mathbb{P}^{\beta}_{g'}(X_k = g', \tau > k, (X_{k+n})_{n=0}^{r} = (x_n)_{n=0}^{r})$$
$$= \sum_{k=0}^{+\infty} \mathbb{P}^{\beta}_{g'}(X_k = g', \tau > k) \mathbb{P}^{\beta}_{g'}((X_n)_{n=0}^{r} = (x_n)_{n=0}^{r})$$
$$= \mathbb{E}^{\beta}_{g'} \left(\sum_{k=0}^{\tau} \mathbf{1}(X_k = g') \right) \mathbb{P}^{\beta}_{g'}((X_n)_{n=0}^{\tau'(E\setminus\{g,g'\})} = (x_n)_{n=0}^{r}).$$

Summing these equalities over the admissible values of $(x_0,\dots,x_r)$, we get that

$$\mathbb{E}^{\beta}_{g'} \left(\sum_{k=0}^{\tau} \mathbf{1}(X_k = g') \right) = (\mathbb{P}^{\beta}_{g'}(X_{\tau'(E\setminus\{g,g'\})} = g))^{-1}. \qquad \square$$

In order to state a large deviation principle for the exit path, let us introduce a few notations. For any $A \subsetneq E$, any $\sigma, \sigma' \in A$, let

$$W_A(\sigma,\sigma') = -\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{E}_\sigma^\beta \left(\sum_{n=0}^{\tau(A)} \mathbf{1}(X_n = \sigma') \right)$$

be the local virtual energy function (the existence of the limit is well known, see Freidlin and Wentzell, 1984). Let us also put by convention for any $\sigma, \sigma' \in E$

$$W_\emptyset(\sigma,\sigma') = 0.$$

Let

$$C_A(\sigma,\sigma') = (W_A(\sigma,\sigma) + V(\sigma,\sigma'))_+, \quad \sigma \in A, \sigma' \in E,$$

$$S_n = \{X_k; 1 \leqslant k \leqslant n\}, \quad \text{with } S_0 = \emptyset,$$

where we have used the notation $(v)_+ = \max\{v, 0\}$. Let us also put for short $\tau' = \tau'(E \setminus \{g, g'\})$.

For any trajectory $\omega \in E^{\mathbb{N}}$, let us introduce the rate function

$$R_A(\omega) = \sum_{n=1}^{\tau'(A)} C_{S_{n-1}(\omega)}(X_{n-1}(\omega), X_n(\omega)), \tag{3}$$

and the stopping times:

$$v_0^A = 0$$
$$v_k^A = \inf\{n; v_{k-1}^A < n \leqslant \tau'(A), C_{S_{n-1}}(X_{n-1}, X_n) > 0 \text{ or } S_n \neq S_{n-1}\}$$

with the convention that $\inf \emptyset = +\infty$ and $X_{+\infty} = \Delta \notin E$. It is proved in [5] that

Theorem 2.1. *For any event $\mathscr{A}$ measurable with respect to $(X_{v_k^A - 1}, X_{v_k^A})_{k=0}^{+\infty}$*

$$\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{P}_{g'}^\beta(\mathscr{A}) = -\inf\{R_A(\omega); \omega \in \mathscr{A}, X_0(\omega) = g'\}.$$

Remark 2.1. The theorem holds not only for events measurable with respect to $(X_{v_k^A})_{k=0}^{+\infty}$, but also for more refined events measurable only with respect to $(X_{v_k^A - 1}, X_{v_k^A})_{k=0}^{+\infty}$. In other words it is possible to study not only bunch of trajectories going through some succession of critical states, but also bunch of trajectories going through some succession of critical edges. Note also that in most cases $v_k^A - 1 > v_{k-1}^A$ and that $(X_{v_k^A})_{k=0}^{+\infty}$ and $(X_{v_k^A - 1})_{k=0}^{+\infty}$ are very likely to have different supports.

We have also a theorem for local times:

Theorem 2.2. *For any $\sigma \in E \setminus \{g, g'\}$*

$$\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{E}_{g'}^\beta \left(\sum_{n=0}^{\tau'} \mathbf{1}(X_n = \sigma) \,\middle|\, \sigma \in S_{\tau'}, X_{\tau'} = g \right) = -W_{E\setminus\{g,g'\}}(\sigma,\sigma).$$

We see that the explicit computation of the large deviation rates described in the two previous theorems requires the computation of the local virtual energies $W_A(\sigma,\sigma)$, $A \subsetneq E$, $\sigma \in A$. These quantities can be easily derived from the *cycle decomposition* of E. Let us review briefly this notion due to Freidlin and Wentzell. We need to introduce first the elevation function

$$H(\sigma,\sigma') = \min\Bigg\{ \max_{i=1,\dots,r} U(\sigma_{i-1}) + V(\sigma_{i-1},\sigma_i);$$
$$\qquad (\sigma_i)_{i=0}^{r} \in E^{r+1}, \sigma_0 = \sigma, \sigma_r = \sigma', 1 \leqslant r < +\infty \Bigg\}. \tag{4}$$

We recall that the virtual energy U is characterised (up to an additive constant) by the fact that the elevation function built from it is symmetric: for any $\sigma,\sigma' \in E$, $H(\sigma,\sigma') = H(\sigma',\sigma)$ (this property is due to Hajek (1988) and to Trouvé (1994) in its widest generality). Using the symmetry of the elevation function, we can introduce the family of equivalence relations $(\mathscr{R}_\lambda)_{\lambda\in\mathbb{R}}$ defined by

$$\mathscr{R}_\lambda = \{(\sigma,\sigma') \in E^2 \mid H(\sigma,\sigma') \leqslant \lambda\} \cup \{(\sigma,\sigma) \in E^2\}.$$

The decomposition of (E,V) into cycles is the family of sets $\mathscr{C}(E,V) = \bigcup_{\lambda\in\mathbb{R}} E/\mathscr{R}_\lambda$. It includes all the one-point sets and forms a tree for the set inclusion relation. For any cycle $C \in \mathscr{C}(E,V)$ we can define the set of ground states and the height of C by

$$G(C) = \{\sigma \in C;\ U(\sigma) = U(C)\}$$

$$\begin{aligned} H(C) &= \inf\{U(\sigma) + V(\sigma,\sigma'); \sigma \in C, \sigma' \notin C\} - U(C) \\ &= \sup\{\lambda;\ C \in E/\mathscr{R}_\lambda\} - U(C). \end{aligned}$$

A very important notion to understand the local behaviour of $(X_n)_{n\in\mathbb{N}}$ in the domain $A \subsetneq E$ is the *maximal partition* of A into cycles

$$\mathscr{M}(A) = \max\{C \in \mathscr{C}(E,V); C \subset A\},$$

where max denotes here the set of maximal elements for the set inclusion relation. Since $(\mathscr{C}(E,V),\subset)$ forms a tree with root E and leaves the one point sets, $\mathscr{M}(A)$ is a partition of A. The local virtual energy in A can be easily expressed in terms of $\mathscr{M}(A)$:

Proposition 2.2. *For any* $\sigma \in C \in \mathscr{M}(A)$

$$W_A(\sigma,\sigma) = U(\sigma) - U(C) - H(C).$$

This proposition means that the exponent describing the time spent in σ by the exit path from A starting at point σ is given by the energy depth at which σ is located into the cycle of the maximal partition of A to which it belongs. This is an intuitive result: the maximal partition of A divides it into "buckets" and the time needed to escape from any of these buckets is exponential in its depth.

The conclusion of this section is that we will be able to compute explicitly the large deviation rate functions relative to the exit path and its local times as soon as we will have computed the virtual energy U and the decomposition into cycles $\mathscr{C}(E,V)$: this is clear from Eq. (3) and Proposition 2.2.

3. The reduction method

A direct computation of the virtual energy U from the matrix tree theorem is in principle possible but in practice often difficult when E is large. The purpose of the reduction method is to express U by solving a succession of *path minimisation* problems, which are more tractable than the graph minimisation problem involved in the direct use of the matrix tree theorem. Another advantage of the method is that it allows to have a partial description of the cycle decomposition, involving only the deepest cycles and the transition probabilities and times between them, even in cases when a description of the full cycle decomposition is intractable, due to the size of the state space and the impossibility to compute the elevation function defined by Eq. (4) in practice for every couple of states.

The idea behind the reduction method is very simple and of probabilistic nature. For any strict subdomain $A \subsetneq E$, let us put

$$\begin{aligned} v_0^A &= 0, \\ v_k^A &= \inf\{n > v_{k-1}^A; X_n \in A\} \end{aligned}$$

and let us consider the reduced (discrete time) process $X_k^A = X_{v_k^A}$.

This reduction mechanism has the following straightforward nice properties:

1. The transitions of X^A satisfies large deviation estimates: there is a large deviation rate function V^A such that for any $\sigma, \sigma' \in A$

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{P}_\sigma^\beta(X_1^A = \sigma') = -V^A(\sigma, \sigma').$$

   This is proved by the identity

$$\mathbb{P}_\sigma^\beta(X_1^A = \sigma') = \mathbb{P}_\sigma^\beta(X_1 = \sigma') + \sum_{\sigma'' \in E \setminus A} \mathbb{P}_\sigma^\beta(X_1 = \sigma'') \mathbb{P}_{\sigma''}^\beta(X_{\tau(E \setminus A)} = \sigma')$$

   combined with the lemma of Freidlin and Wentzell [11] giving a large deviation estimate for exit points. We will note all the classical quantities pertaining to X^A (virtual energy, depth of cycles, etc.) with the superscript A.
2. If $B \subset A$, then $(X^A)^B = X^B$,
3. The invariant probability measure of X^A is the conditional measure $\mu_\beta^A = \mu_\beta(\cdot \mid A)$.
4. When $D \subset A$ the local time of X in $D \cup (E \setminus A)$ is the same as the local time of X^A in D: for any $\sigma, \sigma' \in D$,

$$\mathbb{E}_\sigma^\beta \left(\sum_{n=0}^{\tau(D \cup (E \setminus A))} \mathbf{1}(X_n = \sigma') \right) = \mathbb{E}_\sigma^\beta \left(\sum_{n=0}^{\tau^A(D)} \mathbf{1}(X_n^A = \sigma') \right).$$

5. There are strong links between $\mathscr{C}(E, V)$ and the reduced cycle decomposition $\mathscr{C}(A, V^A)$:

- If $C \in \mathscr{C}(E)$, then $C \cap A \in \mathscr{C}(A)$.

- If $C' \in \mathscr{C}(A)$ and
$$C = \min\{C'' \in \mathscr{C}(E) \mid C' \subset C''\}$$
then $C \cap A = C'$.
- If $C' \in \mathscr{C}(A)$ and
$$C = \max\{C \in \mathscr{C}(E); C \cap A = C'\}$$
then
$$H(C) + U(C) = H^A(C') + U^A(C') + \min_{\sigma \in A} U(\sigma). \tag{5}$$
Consequently, if $G(C) \cap A \neq \emptyset$, then $H^A(C') = H(C)$.

Proof. From Proposition 2.2
$$W_{E\setminus\{\sigma'\}}(\sigma,\sigma) = U(\sigma) - H(C) - U(C),$$
where $C \in \mathscr{M}(E \setminus \{\sigma'\})$ is the cycle containing σ. As the cycle C is the largest cycle containing σ and not containing σ', we have exactly $H(C) + U(C) = H(\sigma,\sigma')$, and
$$W_{E\setminus\{\sigma'\}}(\sigma,\sigma) = U(\sigma) - H(\sigma,\sigma').$$
But in the case when σ and σ' belong to A, then, from the identity of local times and reduced local times
$$W_{E\setminus\{\sigma'\}}(\sigma,\sigma) = W^A_{A\setminus\{\sigma'\}}(\sigma,\sigma),$$
and therefore we have the following relation between the elevations:
$$\begin{aligned} H(\sigma,\sigma') - U(\sigma) &= H^A(\sigma,\sigma') - U^A(\sigma) \\ &= H^A(\sigma,\sigma') - U(\sigma) + \min_{\sigma'' \in A} U(\sigma''). \end{aligned}$$
We have proved that

Lemma 3.1. $H^A(\sigma,\sigma') = H(\sigma,\sigma') - \min_{\sigma'' \in A} U(\sigma'')$.

The links between $\mathscr{C}(E, V)$ and $\mathscr{C}(A, V^A)$ are easy consequences of this lemma.

From this study of cycles, we can deduce a useful result concerning mean exit times:

Proposition 3.1. *For any strict subdomains* $D \subsetneq A \subsetneq E$, *any height* $H > 0$, *as soon as, for any cycle* $C \in \mathscr{M}(D \cup (E \setminus A))$ *such that* $H(C) > H$, *we have* $G(C) \cap A \neq \emptyset$, *then for any* $\sigma \in D$ *such that*
$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}^{\beta}_{\sigma} \tau^A(D) \geqslant H$$
it is true that
$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}^{\beta}_{\sigma}(\tau(D \cup (E \setminus A))) = \lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}^{\beta}_{\sigma}(\tau^A(D)).$$

Loosely speaking, Proposition 3.1 says that we can compute log equivalents of exit times on the reduced chain X^A as soon as A contains some of the states in which the exit path spends most of its time. More precisely, if A intersects the ground states of deep cycles, then the exit times of the reduced chain are of the same order as the exit times of the original chain as soon as they are large enough.

Proof. Let us first remark that

$$\mathbb{E}_\sigma^\beta(\tau^A(D)) = \mathbb{E}_\sigma^\beta\left(\sum_{\sigma'\in D}\sum_{n=0}^{\tau(D\cup(E\setminus A))} \mathbf{1}(X_n=\sigma')\right) \leqslant \mathbb{E}_\sigma^\beta(\tau(D\cup(E\setminus A))).$$

Therefore in the case when

$$\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta(\tau(D\cup(E\setminus A)))\leqslant H$$

the result is obvious. In the other case, let us notice that

$$\begin{aligned}
&\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta(\tau(D\cup(E\setminus A)))\\
&\quad=\sup_{\sigma'\in D\cup(E\setminus A)}\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\sigma')\right)\\
&\quad=\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\hat\sigma)\right)\\
&\quad\leqslant -W_{D\cup(E\setminus A)}(\hat\sigma,\hat\sigma)
\end{aligned}$$

for some $\hat\sigma\in D\cup(E\setminus A)$. Moreover this $\hat\sigma$ whose mean local time is of the order of $\mathbb{E}_\sigma^\beta\tau(D\cup(E\setminus A))$ is such that $-W_{D\cup(E\setminus A)}(\hat\sigma,\hat\sigma)>H$. Now if we define $\hat C$ as the (only) cycle satisfying $\hat\sigma\in\hat C\in\mathscr{M}(D\cup(E\setminus A))$, we have $H(\hat C)\geqslant -U(\hat\sigma)+U(\hat C)+H(\hat C)=-W_{D\cup(E\setminus A)}(\hat\sigma,\hat\sigma)>H$. Therefore there is $\sigma'\in A\cap G(\hat C)=D\cap G(\hat C)$ (because $G(\hat C)\subset D\cup(E\setminus A)$).

We will prove now that the time spent in σ' is (at least) of the order of the time spent in $\hat\sigma$ and therefore of the order of the expectation of the exit time $\tau(D\cup(E\setminus A))$ of the original Markov chain. To do this we will justify the second inequality of the following otherwise obvious chain of equations:

$$\begin{aligned}
\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta(\tau^A(D)) &\geqslant \lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\sigma')\right)\\
&\geqslant \lim_{\beta\to+\infty}\frac{1}{\beta}\log\Big(\mathbb{P}_\sigma^\beta(X_{\tau((D\cup(E\setminus A))\setminus\{\hat\sigma\})}=\hat\sigma)\\
&\qquad\times\mathbb{E}_{\hat\sigma}^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\hat\sigma)\right)\Big)\\
&=\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\hat\sigma)\right)\\
&=\lim_{\beta\to+\infty}\frac{1}{\beta}\log\mathbb{E}_\sigma^\beta(\tau(D\cup(E\setminus A)),
\end{aligned}$$

which proves that

$$\lim_{\beta\to+\infty}\frac{1}{\beta}\log\,\mathbb{E}_\sigma^\beta(\tau^A(D)) = \lim_{\beta\to+\infty}\frac{1}{\beta}\log\,\mathbb{E}_\sigma^\beta(\tau(D\cup(E\setminus A))).$$

Indeed

$$\mathbb{E}_\sigma^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\sigma')\right) \geqslant \mathbb{P}_\sigma^\beta(X_{\tau(D\cup(E\setminus A)\setminus\{\hat\sigma\})}=\hat\sigma)$$

$$\times\mathbb{P}_{\hat\sigma}^\beta(X_{\tau(\hat C\setminus\{\sigma'\})}=\sigma')\mathbb{E}_{\sigma'}^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\sigma')\right).$$

We conclude by observing that, from a well known property of cycles (which may serve as a definition, see Catoni and Cerf, 1997)

$$\lim_{\beta\to+\infty}\frac{1}{\beta}\log\,\mathbb{P}_{\hat\sigma}^\beta(X_{\tau(\hat C\setminus\{\sigma'\})}=\sigma')=0,$$

and that

$$\begin{aligned}
&\lim_{\beta\to+\infty}\frac{1}{\beta}\log\,\mathbb{E}_{\sigma'}^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\sigma')\right)\\
&\quad=H(\hat C)+U(\hat C)-U(\sigma')\\
&\quad\geqslant H(\hat C)+U(\hat C)-U(\hat\sigma)\\
&\quad=\lim_{\beta\to+\infty}\frac{1}{\beta}\log\,\mathbb{E}_{\hat\sigma}^\beta\left(\sum_{n=0}^{\tau(D\cup(E\setminus A))}\mathbf{1}(X_n=\hat\sigma)\right). \qquad\square
\end{aligned}$$

4. The reduction algorithm to compute the virtual energy

From the previous properties of the reduced processes, we can derive the following algorithm:

Build a sequence of domains $A_{|E|}=E\supset A_{|E|-1}\supset\cdots\supset A_1=\{g_1\}$ according to the following backward induction rule:

1. Choose $g_k\in A_k$ such that

$$H^{A_k}(\{g_k\})=\min\{H^{A_k}(\{\sigma\});\sigma\in A_k\} \tag{6}$$

and put $A_{k-1}=A_k\setminus\{g_k\}$.

(The main reason of this choice is contained in the first point of the next proposition: we remove the "shallowest state", and therefore are sure to keep all the ground states of all other remaining cycles. In this procedure, we are sure that in each cycle of the original state space E, the last point to be erased will be a ground state. Removing in one step all the states satisfying (6) would not produce the desired effect (see the proof of the first point of the next proposition: it would be possible to erase g_k and g_k' at the same time, and thus to erase all the ground states in C_k at once). Anyhow it is possible to include many points in one step, and even points with different depths, as will be detailed below under the section "fast reduction algorithm".)

2. Compute

$$V^{A_{k-1}}(\sigma,\sigma') = \min\{V^{A_k}(\sigma,\sigma'),\\ V^{A_k}(\sigma,g_k) + V^{A_k}(g_k,\sigma') - H^{A_k}(\{g_k\})\}, \quad \sigma,\sigma' \in A_{k-1},$$

which is a straightforward consequence of

$$\mathbb{P}^{\beta}_{\sigma}(X_1^{A_{k-1}} = \sigma') = \mathbb{P}^{\beta}_{\sigma}(X_1^{A_k} = \sigma') + \mathbb{P}^{\beta}_{\sigma}(X_1^{A_k} = g_k)\, \mathbb{P}^{\beta}_{g_k}(X^{A_k}_{\tau^{A_k}(\{g_k\})} = \sigma').$$

This construction satisfies the following properties:

Proposition 4.1.

- *For any cycle* $C_k \in \mathscr{C}(A_k)$ *such* that $C_k \cap A_{k-1} \neq \emptyset$ (*i.e.* $C_k \neq \{g_k\}$) *we have that* $G(C_k) \cap A_{k-1} \neq \emptyset$ *and consequently that* $G(C_k \cap A_{k-1}) = G(C_k) \cap A_{k-1}$. *In plain words: in any cycle, the last state to be removed by the reduction algorithm is a ground state.*
- *Therefore for any cycle* $C_k \in \mathscr{C}(A_k)$, *there is* $C \in \mathscr{C}(E)$ *such that* $C \cap A_k = C_k, U(C) = U(C_k)$ *and* $H(C) = H^{A_k}(C_k)$.
- *For any cycle* $C \in \mathscr{C}(E)$, *such that* $H(C) > H^{A_k}(\{g_k\})$ *for some* $k > 1$, *we have* $G(C) \cap A_{k-1} \neq \emptyset$.
- *The virtual energy* U *can be computed in the following inductive way*:

$$U(g_1) = 0,$$
$$U(g_k) = \min\{U(\sigma) + V^{A_k}(\sigma, g_k); \sigma \in A_{k-1}\}\\ -\min\{V^{A_k}(g_k,\sigma); \sigma \in A_{k-1}\}.$$

- *The state* g_2 *is a metastable state*:

$$\max\{H(C); C \in \mathscr{M}(E \setminus \{g_1\})\} = V^{A_2}(g_2, g_1)\\ = H(g_2,g_1) - U(g_2) = \lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}^{\beta}_{g_2}(\tau(E \setminus \{g_1\})).$$

Remark 4.1. Proposition 4.1 does not assume that the situation is not degenerate, in case argmin$U$ is not unique g_2 may (or may not, depending on the shape of the energy landscape) be also a ground state, but it is always one of the states from which getting to g_1 takes the longest time. In case there are several possible pairs (g_2, g_1) the one that is picked up by the reduction algorithm depends on the way the ties are broken. If the ties are broken randomly, every pair will have a positive (although not necessarily uniform) probability to be selected. Moreover, using the terminology of Catoni (1998), we can see that in degenerate cases when there is more than one global ground state, $V^{A_2}(g_2, g_1)$ is equal to the second critical depth (that governs the time after which all the global ground states have been visited when the initial state of the chain is "the worst") that may be larger than the first critical depth (that governs the time after which at least one ground state has been reached, again when the starting point is the most unlucky one).

Proof.

- Let $\tilde{C}_k \in \mathscr{C}(A_k)$ be the smallest cycle such that $\{g_k\} \subsetneq \tilde{C}_k$ and let us consider some $C'_k \in \mathscr{M}(\tilde{C}_k \setminus \{g_k\})$ and some $g'_k \in G(C'_k)$. As $H^{A_k}(g_k, g'_k) = H^{A_k}(g'_k, g_k)$, we have

$$\begin{aligned} U(g_k) + H^{A_k}(\{g_k\}) &= U(g'_k) + H^{A_k}(C'_k) \\ &\geqslant U(g'_k) + H^{A_k}(\{g'_k\}) \\ &\geqslant U(g'_k) + H^{A_k}(\{g_k\}). \end{aligned}$$

  Therefore $U(g_k) \geqslant U(g'_k)$. Now for any given cycle $C_k \in \mathscr{C}(A_k)$ such that $C_k \cap A_{k-1} \neq \emptyset$ and $g_k \in G(C_k)$ we have $\tilde{C}_k \subset C_k$ and thus $g'_k \in G(C_k) \setminus \{g_k\} = G(C_k) \cap A_{k-1} \neq \emptyset$. (Note that $g'_k \in G(C_k)$ because $g'_k \in C_k$, $g_k \in G(C_k)$ and $U(g'_k) \leqslant U(g_k)$, and therefore necessarily $U(g'_k) = U(g_k) = \min_{\sigma \in C_k} U(\sigma)$.)
- Let $C \in \mathscr{C}(E)$ be the largest one such that $C \cap A_k = C_k$, then, according to Eq. (5) (written here in terms of U only), $H(C) + U(C) = H^{A_k}(C_k) + U(C_k)$. Applying successively the first point to $C \cap A_j \in \mathscr{C}(A_j)$ for $j = |E|, \ldots, k+1$ we see that $G(C) \cap A_k \neq \emptyset$ and so that $G(C \cap A_k) = G(C) \cap A_k$. Thus $U(C) = U(C_k)$ and $H(C) = H^{A_k}(C_k)$.
- Take $C \in \mathscr{C}(E)$ such that $H(C) > H^{A_k}(\{g_k\})$. For any j such that $k \leqslant j \leqslant |E|$, we have

$$H^{A_j}(\{g_j\}) \geqslant H^{A_{j+1}}(\{g_j\}) \geqslant H^{A_{j+1}}(\{g_{j+1}\}),$$

  and therefore $H(C) > H^{A_k}(\{g_k\}) \geqslant H^{A_j}(\{g_j\})$. Assume now that $G(C) \cap A_j \neq \emptyset$, which is certainly true when $j = |E|$, then $H^{A_j}(C \cap A_j) \geqslant H(C) > H^{A_j}(\{g_j\})$, therefore $C \cap A_{j-1} \neq \emptyset$ thus $G(C) \cap A_{j-1} \neq \emptyset$ and $G(C \cap A_{j-1}) = G(C) \cap A_{j-1}$ from the first point. Therefore by induction on j, $G(C) \cap A_{k-1} \neq \emptyset$.
- A special case of the previous point is that $G(E) \cap A_1 \neq \emptyset$, which means that $U(g_1) = 0$. The expression given for $U(g_k)$ comes from the invariance of $\mu_\beta(\cdot \mid A_k)$:

$$\mu_\beta(\{g_k\} \mid A_k) \sum_{\sigma \in A_{k-1}} p_\beta^{A_k}(g_k, \sigma) = \sum_{\sigma \in A_{k-1}} \mu_\beta(\{\sigma\} \mid A_k) p_\beta^{A_k}(\sigma, g_k).$$

- Let $C \in \mathscr{C}(E)$ be such that $H(C) > H^{A_2}(\{g_2\}) = V^{A_2}(g_2, g_1)$. Then from the third point $G(C) \cap \{g_1\} \neq \emptyset$. This proves that

$$\max\{H(C) \colon C \in \mathscr{M}(E \setminus \{g_1\})\} \leqslant V^{A_2}(g_2, g_1).$$

  To prove equality, we apply the second point to $\{g_2\} \in \mathscr{C}(A_2)$, to find a cycle $C \subset E \setminus \{g_1\}$ such that $H(C) = V^{A_2}(g_2, g_1)$. Thus the hypotheses of Proposition 2 are fulfilled for $H = V^{A_2}(g_2, g_1)$, $A = A_2$, $D = \{g_2\}$, as $D \cup (E \setminus A) = E \setminus \{g_1\}$, this proves that

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}_{g_2}^{\beta}(\tau(E \setminus \{g_1\})) = \lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}_{g_2}^{\beta} \tau^{\{g_1, g_2\}}(\{g_2\}) = V^{A_2}(g_2, g_1).$$

  Eventually,

$$V^{A_2}(g_2, g_1) = H^{A_2}(g_2, g_1) - U^{A_2}(g_2) = H(g_2, g_1) - U(g_2),$$

  as a consequence of Lemma 3.1. This achieves to prove the last point. $\square$

5. The fast reduction algorithm

The aim of this section is to describe a fast reduction algorithm in which several steps of the reduction algorithm are performed in one stage when this is possible. In the (slow) reduction algorithm, the computation of the reduced rate function V^{A_k} from $V^{A_{k+1}}$ uses paths of length 1 or 2. Here we need to introduce broader families of "reduction paths".

For any path $\gamma = (\gamma_i)_{i=0}^r$ we can define the reduced path γ^A by

$$\begin{aligned} v^A(0) &= \inf\{i; \gamma_i \in A\},\\ v^A(s) &= \inf\{i > v^A(s-1); \gamma_i \in A\},\\ r^A &= \inf\{s; v^A(s+1) = +\infty\},\\ \gamma_s^A &= \gamma_{v^A(s)}, \quad s = 0, \ldots, r^A(\gamma). \end{aligned}$$

In the following, we let $(A_k)_{k=1}^{|E|}$ be the sequence of subdomains corresponding to some (slow) reduction algorithm. We first need to establish a formula to compute directly V^{A_k} from V^{A_j} for any $k < j$.

For this purpose, we define the set of reduction paths $\mathscr{V}_k^j$ of A_j to A_k by

$$\begin{aligned} \mathscr{V}_k^j = \{&(\gamma_i)_{i=0}^r \in E^{r+1};\ r \geqslant 1, \{\gamma_0, \gamma_r\} \subset A_k,\ (\gamma_i)_{i=1}^{r-1} \in (A_j \setminus A_k)^{r-1},\\ &\gamma_{i-1}^{A_m \setminus A_k} \neq \gamma_i^{A_m \setminus A_k},\ k < m \leqslant j,\ 0 < i \leqslant r^{A_m \setminus A_k}(\gamma)\} \end{aligned}$$

(no reduced path $\gamma^{A_m \setminus A_k}$ should contain consecutive repetitions, but γ_0 can be the same as γ_r).

To understand the next proposition, the reader should take some time to get familiar with this notion. Here are some examples:

$$\begin{aligned} &(g_1, g_3, g_2, g_3, g_1) \in \mathscr{V}_1^j, \quad \text{for any } j \geqslant 3,\\ &(g_1, g_4, g_3, g_4, g_5, g_2) \in \mathscr{V}_2^j, \quad \text{for any } j \geqslant 5, \end{aligned}$$

but

$$(g_1, g_4, g_5, g_4, g_3, g_2) \notin \mathscr{V}_2^j, \quad \text{for any } j \geqslant 5,$$

because g_4 appears at two consecutive locations in the reduction of this last path to A_4.

Note that the set $\mathscr{V}_k^j$ is finite, because in any of its paths, g_{k+1} can appear at most once, g_{k+2} at most twice, and more generally $g_{k+\ell}$ at most $2^{\ell-1}$ times.

We also need to define the depth $\tilde{H}(\sigma)$ of any state σ of E. Since $E = \{g_k;\ k = 1, \ldots, |E|\}$, we can do this putting $\tilde{H}(g_k) = H^{A_k}(\{g_k\})$, with the convention that $H^{A_1}(\{g_1\}) = +\infty$. Note that in degenerate situations the (slow) reduction algorithm is not unique, and that the definition of $\tilde{H}(\sigma)$ depends on the choice of some (arbitrary) reduction algorithm.

Proposition 5.1. *For any* j, $k < j \leqslant |E|$, *any* $\sigma, \sigma' \in A_k$,

$$V^{A_k}(\sigma,\sigma') = \min\left\{ V^{A_j}(\gamma_0,\gamma_1) + \sum_{i=2}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - \tilde{H}(\gamma_{i-1});\right.$$
$$\left. r\geqslant 1,\ \gamma_0=\sigma,\ \gamma_r=\sigma',\ (\gamma_i)_{i=0}^r \in \mathscr{V}_k^j \right\}.$$

Proof. By induction on $j = k+1,\ldots,|E|$. The formula is true for $j=k+1$, because it reduces to the relation between V^{A_k} and $V^{A_{k+1}}$ established on the occasion of the (slow) reduction algorithm. Assume that the result holds for $j-1$. Let $(\gamma_i)_{i=0}^r \in \mathscr{V}_k^j$ be such that $\gamma_0=\sigma$, $\gamma_r=\sigma'$. Let $v = v^{A_{j-1}}$ be the change of time corresponding to the reduction to A_{j-1} of γ. Then $v(i) \in \{v(i-1)+1, v(i-1)+2\}$ and in the case when $v(i) = v(i-1)+2$, $\gamma_{v(i-1)+1} = g_j$, because $\{g_j\} = A_j \setminus A_{j-1}$ and because repetitions are forbidden in $\mathscr{V}_k^j$. Thus

$$V^{A_{j-1}}(\gamma_{v(i-1)},\gamma_{v(i)}) \leqslant V^{A_j}(\gamma_{v(i-1)},\gamma_{v(i-1)+1}) + \sum_{s=v(i-1)+2}^{v(i)} V^{A_j}(\gamma_{s-1},\gamma_s) - \tilde{H}(\gamma_{s-1}),$$

where the sum is assumed to be equal to 0 when the summation is between $s = v(i-1)+2$ and $s = v(i-1)+1$, that is when $v(i) = v(i-1)+1$. From the induction assumption we have

$$V^{A_k}(\sigma,\sigma') \leqslant V^{A_{j-1}}(\gamma_{v(0)},\gamma_{v(1)}) + \sum_{i=2}^{r^{A_{j-1}}} V^{A_{j-1}}(\gamma_{v(i-1)},\gamma_{v(i)}) - \tilde{H}(\gamma_{v(i-1)}).$$

Substituting the former inequality into the right-hand side of the latter, we get that

$$V^{A_k}(\sigma,\sigma') \leqslant V^{A_j}(\gamma_0,\gamma_1) + \sum_{i=2}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - \tilde{H}(\gamma_{i-1}).$$

On the other hand we can build $(\psi_i)_{i=0}^s \in \mathscr{V}_k^{j-1}$, $\psi_0=\sigma$, $\psi_s=\sigma'$, such that

$$V^{A_k}(\sigma,\sigma') = V^{A_{j-1}}(\psi_0,\psi_1) + \sum_{i=2}^{s} V^{A_{j-1}}(\psi_{i-1},\psi_i) - \tilde{H}(\psi_{i-1}),$$

and using the identity

$$V^{A_{j-1}}(\psi_{i-1},\psi_i) = \min\{V^{A_j}(\psi_{i-1},\psi_i), V^{A_j}(\psi_{i-1},g_j) + V^{A_j}(g_j,\psi_i) - \tilde{H}(g_j)\},$$

we can build $(\gamma_i)_{i=0}^r \in \mathscr{V}_k^j$ such that $(\gamma_i^{A_{j-1}}) = (\psi_i)$ and

$$V^{A_k}(\sigma,\sigma') = V^{A_j}(\gamma_0,\gamma_1) + \sum_{i=2}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - \tilde{H}(\gamma_{i-1}). \qquad \square$$

Let us deal now with the computation of energies. For any $1 \leqslant k < j \leqslant |E|$, let us consider the set $\mathscr{U}_k^j$ of reduction paths defined by

$$\mathscr{U}_k^j = \{(\gamma_i)_{i=0}^r;\ r\geqslant 1,\ \gamma_0 \in A_k,\ (\gamma_i)_{i=1}^r \in (A_j\setminus A_k)^r,\ \gamma_i^{A_m} \neq \gamma_{i-1}^{A_m},$$
$$k < m \leqslant j,\ 1\leqslant i \leqslant r^{A_m}(\gamma)\}.$$

Then

Proposition 5.2. *For any* $\sigma' \in (A_j \setminus A_k)$

$$U(\sigma') = \min\left\{ U(\gamma_0) + \sum_{i=1}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - \tilde{H}(\gamma_i);\ r \geqslant 1,\ \gamma_r = \sigma',\ (\gamma_i)_{i=0}^{r} \in \mathscr{U}_k^j \right\}.$$

Proof. We have $\sigma' = g_s$ for some s such that $k < s \leqslant j$. Let us assume by induction that the proposition has been proved for any $\sigma' \in A_{s-1} \setminus A_k$, then we have

$$\begin{aligned} U(g_s) + \tilde{H}(g_s) &= \min\{U(\sigma) + V^{A_s}(\sigma, g_s);\ \sigma \in A_{s-1}\},\\ &= \min\Bigg\{ U(\gamma_0) + \sum_{i=1}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - \tilde{H}(\gamma_i)\\ &\qquad + V^{A_j}(\gamma'_0,\gamma'_1) + \sum_{\ell=2}^{r'} V^{A_j}(\gamma'_{\ell-1},\gamma'_\ell) - \tilde{H}(\gamma'_{\ell-1});\\ &\qquad r \geqslant 1, (\gamma_i)_{i=0}^{r} \in \mathscr{U}_k^j, \gamma_r \in A_{s-1},\\ &\qquad r' \geqslant 1, (\gamma'_\ell)_{\ell=0}^{r'} \in \mathscr{V}_s^j, \gamma'_0 = \gamma_r, \gamma'_{r'} = g_s \Bigg\}\\ &= \min\Bigg\{ U(\gamma_0) + \sum_{i=1}^{r''} V^{A_j}(\gamma''_{i-1},\gamma''_i) - \tilde{H}(\gamma''_i) + \tilde{H}(g_s),\\ &\qquad r'' \geqslant 1,\ (\gamma''_i)_{i=0}^{r''} \in \mathscr{U}_k^j,\ \gamma''_{r''} = g_s \Bigg\}, \end{aligned}$$

because any path $(\gamma''_i)_{i=0}^{r''} \in \mathscr{U}_k^j$, $\gamma''_{r''} = g_s$, can be decomposed into the concatenation of $(\gamma_i)_{i=1}^{r} \in \mathscr{U}_k^j$, $\gamma_r \in A_{s-1}$, and $(\gamma'_i)_{i=0}^{r'} \in \mathscr{V}_s^j$, $\gamma'_0 = \gamma_r$, $\gamma'_{r'} = g_s$, by considering the last index i for which $\gamma''_i \in A_{s-1}$. $\square$

Due to the two previous propositions, in the case when we can compute $\tilde{H}(\sigma)$ for $\sigma \in (A_j \setminus A_k)$ directly from V^{A_j}, we can jump in one step from A_j to A_k and back in the two passes of the reduction algorithm.

This will be the case when $\mathscr{M}^{A_j}(A_j \setminus A_k) = \{\{g\};\ g \in A_j \setminus A_k\}$ or equivalently when for any $\sigma \in A_j \setminus A_k$, there is a path $(\sigma_i)_{i=0}^{r}$ such that $r \geqslant 1$, $(\sigma_i)_{i=0}^{r-1} \in (A_j \setminus A_k)^r$, $\sigma_0 = \sigma$, $\sigma_r \in A_k$ and $V^{A_j}(\sigma_{i-1},\sigma_i) = H^{A_j}(\{\sigma_{i-1}\})$.

Proposition 5.3. *Assume that we have performed a reduction algorithm from the initial step $A_{|E|} = E$ to the reduction step A_j, and assume that for some subset A_k of A_j of size $|A_k| = k$, we have $\mathscr{M}^{A_j}(A_j \setminus A_k) = \{\{\sigma\};\ \sigma \in A_j \setminus A_k\}$, and*

$$\max\{H^{A_j}(\{\sigma\});\ \sigma \in A_j \setminus A_k\} \leqslant \min\{H^{A_j}(\{\sigma\});\ \sigma \in A_k\},$$

then for any sequence $(g_s)_{s=k+1}^{j} \in (A_j \setminus A_k)^{j-k}$ such that $H^{A_j}(g_s)$ is non decreasing, the sequence of reduction sets $A_i = A_j \setminus \{g_s;\ i < s \leqslant j\}$, $k \leqslant i < j$, is part of some valid (slow) reduction algorithm, $\tilde{H}(g_s) = H^{A_j}(\{g_s\})$, $k < s \leqslant j$, and the constraints $\mathscr{V}_k^j$ and

$\mathscr{U}_k^j$ on the reduction paths can be relaxed to

$$V^{A_k}(\sigma,\sigma') = \min\Bigg\{ V^{A_j}(\gamma_0,\gamma_1) + \sum_{i=2}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - H^{A_j}(\{\gamma_{i-1}\}); \\ r \geqslant 1,\ \gamma_0=\sigma,\ \gamma_r=\sigma',\ (\gamma_i)_{i=1}^{r-1} \in (A_j\setminus A_k)^{r-1}, \\ \gamma_i \neq \gamma_{i-1},\ i=1,\dots,r \Bigg\},\quad \sigma,\sigma'\in A_k,$$

and

$$U(\sigma') = \min\Bigg\{ U(\gamma_0) + \sum_{i=1}^{r} V^{A_j}(\gamma_{i-1},\gamma_i) - H^{A_j}(\{\gamma_i\}); \\ r \geqslant 1,\ \gamma_0 \in A_k,\ \gamma_r=\sigma',\ (\gamma_i)_{i=1}^{r-1} \in (A_j\setminus A_k)^{r-1}, \\ \gamma_i \neq \gamma_{i-1},\ i=1,\dots,r \Bigg\},\quad \sigma'\in A_j\setminus A_k.$$

Proof. Let us establish first by decreasing induction that A_i, $i=j-1,\dots,k$, is a valid sequence of reduction sets. This is straightforward from the definition of the (slow) reduction algorithm when $i=j-1$. Assume that we have proved that $\{A_s;\ s=j-1,\dots,i\}$ is valid and that for any s, $i<s\leqslant j$, $\tilde H(g_s) \stackrel{\text{def}}{=} H^{A_s}(\{g_s\}) = H^{A_j}(\{g_s\})$. It is enough to prove that $H^{A_i}(\{\sigma\}) = H^{A_j}(\{\sigma\})$ for any $\sigma\in A_i$ to make one induction step. We have

$$\begin{aligned} H^{A_i}(\{\sigma\}) &= \min\{V^{A_i}(\sigma,\sigma');\ \sigma'\in A_i\setminus\{\sigma\}\} \\ &= \min\Bigg\{ V^{A_j}(\gamma_0,\gamma_1) + \sum_{s=2}^{r} V^{A_j}(\gamma_{s-1},\gamma_s) - \tilde H(\gamma_{s-1}); \\ &\qquad r\geqslant 1,\ \gamma_0=\sigma,\ \gamma_r\in A_i\setminus\{\sigma\},\ (\gamma_s)_{s=0}^r \in \mathscr{V}_i^j \Bigg\} \\ &= \min\Bigg\{ V^{A_j}(\gamma_0,\gamma_1) + \sum_{s=2}^{r} V^{A_j}(\gamma_{s-1},\gamma_s) - H^{A_j}(\gamma_{s-1}); \\ &\qquad r\geqslant 1,\ \gamma_0=\sigma,\ \gamma_r\in A_i\setminus\{\sigma\},\ (\gamma_s)_{s=0}^r \in \mathscr{V}_i^j \Bigg\} \\ &= \min\{V^{A_j}(\gamma_0,\gamma_1);\ r\geqslant 1,\ \gamma_0=\sigma,\ \gamma_r\in A_i\setminus\{\sigma\},\ (\gamma_s)_{s=0}^r\in\mathscr{V}_i^j\} \\ &= \min\{V^{A_j}(\sigma,\sigma');\ \sigma'\in A_i\setminus\{\sigma\}\}. \end{aligned}$$

The constraints $\mathscr{V}_k^j$ and $\mathscr{U}_k^j$ can be weakened as mentioned in the proposition, because all the quantities $V^{A_j}(\gamma_{i-1},\gamma_i) - H^{A_j}(\gamma_{i-1})$ are positive, and therefore any path $(\gamma_i)_{i=0}^r$ such that $\{\gamma_0,\gamma_r\}\in A_k$ (resp. $\gamma_0\in A_k$ and $\gamma_r\in A_j\setminus A_k$) can be reduced to a path in $\mathscr{V}_k^j$ (resp. in $\mathscr{U}_k^j$) of lower cost by pruning all its loops. □

6. The loop erased exit path

So far we have used the reduction method as a tool for computations. In this section we will give a probabilistic interpretation to the minimisation problem appearing in

Proposition 5.1. More precisely we will show that Proposition 5.1 can be interpreted as a large deviation contraction formula linking the large deviation rate for the transition $\mathbb{P}_\sigma^\beta(X_1^{A_k}=\sigma')$ with the large deviation rate for the values of some loop erased excursion path, defined below as some function of the excursion path $(X_i^{A_j})_{i=0}^{\tau'(A_j\setminus A_k)}$.

A special case of interest is of course the study of the loop erased exit path built from $(X_i)_{i=0}^{\tau'(E\setminus A_2)}$.

The first thing we need to do is to define for any $k<j$ a "loop eraser" map Γ_k^j from the possible values of the excursion path of $(X_i)_{i\in\mathbb{N}}$ from $\sigma\in A_k$ to $\sigma'\in A_k$ to the set of reduction paths $\mathscr{V}_k^j$. Namely, $k<j$ and $\sigma,\sigma'\in A_k$ being fixed, we define

$$\Gamma_k^j:\{(x_i)_{i=0}^r\in E^{r+1};\ x_0=\sigma,\ x_r=\sigma',\ (x_i)_{i=1}^{r-1}\in(E\setminus A_k)^{r-1}\}\to\mathscr{V}_k^j$$

by the following induction: let us put

$$\Delta_k(x)=\{0,1\},\quad \theta(0)=0,\quad \tau(1)=r,$$

and assume that we have built the times $\{(\theta(\delta),\tau(\delta));\ \delta\in\Delta_{j-1}(x)\}$ and the function $\{\ell(\delta)\in\mathbb{N};\ \delta\in\Delta_{j-1}(x)\cap\,]0,1[\}$ where $\Delta_{j-1}(x)$ is some finite subset of $[0,1]$. If (δ_-,δ) are two consecutive points in $\Delta_{j-1}(x)$, we define

$$\ell\left(\frac{\delta_-+\delta}{2}\right)=\begin{cases} j & \text{if } g_j\in\{x_i;\ \theta(\delta_-)<i<\tau(\delta)\},\\ +\infty & \text{otherwise.}\end{cases}$$

We let $\Delta_j(x)$ be the union of $\Delta_{j-1}(x)$ and the values $(\delta_-+\delta)/2$ for which $\ell((\delta_-+\delta)//2)$ is equal to j. For these values, we set

$$\tau\left(\frac{\delta_-+\delta}{2}\right)=\inf\{i;\ \theta(\delta_-)<i<\tau(\delta)\text{ and } x_i=g_j\},$$

$$\theta\left(\frac{\delta_-+\delta}{2}\right)=\sup\{i;\ \theta(\delta_-)<i<\tau(\delta)\text{ and } x_i=g_j\}.$$

Then we can sort $\Delta_j(x)$ in increasing order : $\Delta_j(x)=\{\delta_i;\ 0\leqslant i<|\Delta_j(x)|\}$ and put $\gamma_i=g_{\ell(\delta_i)}$; $0<i<|\Delta_j(x)|-1$, $\gamma_0=\sigma$ and $\gamma_{|\Delta_j(x)|-1}=\sigma'$. It is easy to check that $(\gamma_i)_{i=0}^{|\Delta_j(x)|-1}\in\mathscr{V}_k^j$, and we define $\Gamma_k^j((x_0,\dots,x_r))=(\gamma_i)_{i=0}^{|\Delta_j(x)|-1}$.

The loop eraser map has nice reduction properties. More precisely for any $k<j<s$, $\Gamma_k^j(x)=\Gamma_k^s(x)^{A_j}$ and $\Gamma_k^j(x)=\Gamma_k^j(x^{A_j})$. It is also a projection: $\Gamma_k^j(\Gamma_k^j(x))=\Gamma_k^j(x)$, and the set $\Delta_j(x)$ and the function $\{\ell(\delta);\ \delta\in\Delta_j\cap\,]0,1[\}$ depend only on $\Gamma_k^j(x)$ and not on x itself.

The loop eraser map Γ_k^j can also be described more informally in the following way: first go from x to the reduced path x^{A_j}, then erase from x^{A_j} the loops starting at g_{k+1} by deleting all that is between the first and last visit to g_{k+1}, then delete all the loops starting at g_{k+2} that do not hit g_{k+1}... then delete all the loops starting at g_{k+s} that do not hit $A_{k+s-1}\dots$. Note that the loop erased path $\gamma=\Gamma_k^j((x_0,\dots,x_r))$ can still contain some loops, but that it cannot contain too many of them, since it is a reduction path, and the set of reduction paths $\mathscr{V}_k^j$ is finite: indeed g_{k+1} can appear only once in γ, g_{k+2} can appear only twice, and more generally, g_{k+s} can only appear 2^{s-1} times.

It is easy to check by induction that the loop erased excursion path satisfies the following large deviation principle:

Theorem 6.1. *Let us define the cost $R_j(\gamma)$ of the reduction path $\gamma \in \mathscr{V}_k^j$ by*

$$R_j(\gamma) = V^{A_j}(\gamma_0, \gamma_1) + \sum_{i=2}^{r} V^{A_j}(\gamma_{i-1}, \gamma_i) - \tilde{H}(\gamma_{i-1}),$$

then for any subset F of $\mathscr{V}_k^j$

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{P}_\sigma^\beta(\Gamma_k^j((X_i^{A_j})_{i=0}^{\tau'(A_j \setminus A_k)}) \in F) = -\inf_{\gamma \in F} R_j(\gamma).$$

Proof.

- As $\mathscr{V}_k^j$ is a finite set, it is enough to study the case when $F = \{\gamma\}$ is a one point set, so that in all this proof, γ will be some fixed reduction path in $\mathscr{V}_k^j$.
- Let us put for short in this proof $\Gamma_k^j((X^{A_j})_0^{\tau'(A_j \setminus A_k)}) = \Gamma_k^j(X)$.
- We are going to compute $\mathbb{P}_\sigma^\beta(\Gamma_k^j(X) = \gamma)$ as a function of $\mathbb{P}_\sigma^\beta(\Gamma_k^{j-1}(X) = \gamma^{A_{j-1}})$.
- As $\Delta_j(x) = \Delta_j(\Gamma_j^k(x))$, on the set $\Gamma_k^j(X) = \gamma$, $\Delta_j(X)$ is nonrandom and equal to $\Delta_j(\gamma)$. On this set, the function $\{\ell(\delta);\ \delta \in \Delta_j(\gamma)\}$ appearing in the construction of $\Gamma_j^k(X)$ is also a nonrandom object depending only on γ. On the other hand, $\{\tau(\delta), \theta(\delta);\ \delta \in \Delta_j(\gamma)\}$ are random times.
- For any consecutive pair (δ_-, δ) in Δ_{j-1}, we have that $(\delta_- + \delta)/2 \in \Delta_j$ if and only if g_j belongs to the path $(X_{\theta(\delta_-)}, \dots, X_{\tau(\delta)})$. The probability of such an event knowing that $\Gamma_k^{j-1}(X) = \gamma^{A_{j-1}}$ is the probability that $X_1^{A_j} = g_j$ knowing that $X_0^{A_j} = g_{\ell(\delta_-)}$ and that $X_1^{A_{j-1}} = g_{\ell(\delta)}$, that is

$$\frac{\mathbb{P}_{g_{\ell(\delta_-)}}^\beta(X_1^{A_j} = g_j)\mathbb{P}_{g_j}^\beta(X_{\tau^{A_j}(\{g_j\})}^{A_j} = g_{\ell(\delta)})}{\mathbb{P}_{g_{\ell(\delta_-)}}^\beta(X_1^{A_{j-1}} = g_{\ell(\delta)})}.$$

  In the same way the probability that $(\delta_- + \delta)/2 \notin \Delta_j$ knowing $\Gamma_k^{j-1}(X) = \gamma^{A_{j-1}}$ is the probability that g_j does not belong to the sequence $(X_{\theta(\delta_-)}, \dots, X_{\tau(\delta)})$ knowing that $X_{\theta(\delta_-)} = g_{\ell(\delta_-)}$, $X_{\tau(\delta)} = g_{\ell(\delta)}$, and is equal to

$$\frac{\mathbb{P}_{g_{\ell(\delta_-)}}^\beta(X_1^{A_j} = g_{\ell(\delta)})}{\mathbb{P}_{g_{\ell(\delta_-)}}^\beta(X_1^{A_{j-1}} = g_{\ell(\delta)})}.$$

- Moreover, applying repeatedly (a straightforward extension of) Proposition 2.1 at the random times $\theta(\delta)$, $\delta \in \Delta_{j-1}$, and the Markov property at the random times $\tau(\delta)$, $\delta \in \Delta_{j-1}$, we see that these events are conditionally independent knowing that $\Gamma_k^{j-1}(X) = \gamma^{A_{j-1}}$.

- Therefore we have that

$$\mathbb{P}^{\beta}_{\sigma}(\Gamma^{j}_{k}(X)=\gamma)=\mathbb{P}^{\beta}_{\sigma}(\Gamma^{j-1}_{k}(X)=\gamma^{A_{j-1}})$$
$$\times\prod_{\substack{\delta\in\Delta_{j-1}\setminus\{0\},\\(\delta_{-}+\delta)/2\in\Delta_j}}\frac{\mathbb{P}^{\beta}_{g_{\ell(\delta_-)}}(X^{A_j}_1=g_j)\mathbb{P}^{\beta}_{g_j}(X^{A_j}_{\tau^{A_j}(\{g_j\})}=g_{\ell(\delta)})}{\mathbb{P}^{\beta}_{g_{\ell(\delta_-)}}(X^{A_{j-1}}_1=g_{\ell(\delta)})}$$
$$\times\prod_{\substack{\delta\in\Delta_{j-1}\setminus\{0\},\\(\delta_{-}+\delta)/2\notin\Delta_j}}\frac{\mathbb{P}^{\beta}_{g_{\ell(\delta_-)}}(X^{A_j}_1=g_{\ell(\delta)})}{\mathbb{P}^{\beta}_{g_{\ell(\delta_-)}}(X^{A_{j-1}}_1=g_{\ell(\delta)})},$$

  where δ_- is the predecessor of δ in Δ_{j-1}.
- Taking the logarithm of this expression, it is elementary to check the theorem by induction on $j=k+1,\dots,|E|$ for any fixed value of k. $\square$

To make a connection with the exit cycle path studied in Catoni and Cerf (1997), we can notice that if $\pi(\gamma)$ is the cycle path of γ, that is the sequence of maximal cycles of $\mathcal{M}^{A_j}(A_j\setminus A_k)$ visited by γ, and if $\pi(x)$ is the cycle path of $x=(x_0,\dots,x_r)$, then $\pi(\gamma)$ is deduced from $\pi(x)$ by erasing some suitable loops. We can define (see Catoni and Cerf, 1997) the cost of a cycle path $\pi(x)=(x_0,C_0,\dots,C_\ell,x_r)$, where $C_k\in\mathcal{M}(A_j\setminus A_k)$ for $k=0,\dots,\ell$, as

$$R(\pi)=\inf_{y\in C_0}V^{A_j}(x_0,y)+\sum_{k=1}^{\ell}\inf\{(W_{C_{k-1}}(y,y)+V(y,z))_+;\,y\in C_{k-1},z\in C_k\}$$
$$+\inf\{(W_{C_\ell}(y,y)+V(y,x_r))_+;\,y\in C_\ell\}.$$

Then

$$\mathbb{P}^{\beta}_{\sigma}(\Gamma^{j}_{k}((X^{A_j}_i)_{i=0}^{\tau'(A_j\setminus A_k)})=\gamma)\leqslant\mathbb{P}^{\beta}_{\sigma}(\pi(\gamma)\in\text{ Loop-erased}(\pi((X^{A_j}_i)_{i=0}^{\tau'(A_j\setminus A_k)})))$$
$$\leqslant\mathbb{P}^{\beta}_{\sigma}(R(\pi((X_i)_{i=0}^{\tau'(A_j\setminus A_k)}))\geqslant R(\pi(\gamma))),$$

where Loop-erased($\pi(x)$) is the set of cycle paths obtained from $\pi(x)$ by erasing some loops. This is because erasing loops in a cycle path cannot increase its cost. Taking the logarithms, dividing by β and going to the limit, this proves (from the large deviation principle for cycle paths established in Catoni and Cerf, 1997) that

Proposition 6.1. *For any reduction path $\gamma\in\mathcal{V}^{j}_{k}$,*

$$R_j(\gamma)\geqslant R(\pi(\gamma)):$$

the cost of a reduction path is not lower than the cost of its cycle path.

As a corollary of this proposition, if

$$R((\gamma_0,\dots,\gamma_r))=\inf\{R(\gamma'_0,\dots,\gamma'_{r'});\,\gamma'\in\mathcal{V}^{j}_{k},\gamma'_0=\sigma,\gamma'_r=\sigma'\},$$

that is if γ is a reduction path of minimal cost joining σ to σ', then the cycle path $\pi(\gamma)$ is also a cycle path of minimal cost joining σ to σ'. Conversely, if π is a cycle path of minimal cost joining σ to σ', then it has to contain at least one trajectory x such that $R_j(\Gamma^{j}_{k}(x))$ is of minimal cost, and therefore such that $R(\pi(\Gamma^{j}_{k}(x)))$ is of minimal cost. Therefore $\pi = \pi(x)$ is the cycle path of a reduction path of minimal cost with maybe some added loops of null cost.

Another interesting feature of the loop erased excursion is that the map Γ_k^j we use to erase loops does not erase the deepest state visited by the excursion path (the one with the largest value of $\tilde{H}$). This implies that

Proposition 6.2. *For any reduction path* $\gamma = (\gamma_i)_{i=0}^r \in \mathscr{V}_k^j$ *such that* $\gamma_0 = \sigma \in A_k$

$$\lim_{\beta \to +\infty} \frac{1}{\beta} \log \mathbb{E}_\sigma^\beta(\tau'(A_j \setminus A_k) | \Gamma_k^j((X_i^{A_j})_{i=0}^{\tau'(A_j \setminus A_k)}) = (\gamma_i)_{i=0}^r)$$
$$= \sup\{\tilde{H}(\gamma_i); 0 < i < r\}.$$

It is also to be noted that the structure of the loops that have been erased in the erased excursion path can be explored in a hierarchical way, since Theorem 6.1 applies in the case when the reduction path γ is a loop. Namely, the loops that have been erased starting from state g by the loop eraser map Γ_k^j are typically loops of cost no greater than $\tilde{H}(g)$. This can be seen easily by applying the Markov property at point g.

All this discussion applies of course to the study of the metastability, which is nothing but the special case when $k = 2$, $j = |E|$, $\sigma = g_2$ and $\sigma' = g_1$. In brief, the conclusion is that with a probability close to one when β is large the exit path is equal to some reduction path of minimal cost up to some loop removals.

7. Application to the biased majority vote process

The biased majority vote process is a nonreversible continuous time dynamics defined in the following way: we put $E = \{0,1\}^{\mathrm{T}}$ where $T = (\mathbb{Z}/N\mathbb{Z})^2$ and we consider on the torus T the neighbourhood structure $\mathscr{N}(x) = \{y \in T; ||x - y|| = 1\}$ (where $||x - y||$ stands for the Euclidean norm) and the infinitesimal generators

$$(L_\beta f)(\sigma) = \sum_{x \in T} c_\beta(x,\sigma)(f(\sigma^x) - f(\sigma)),$$

where

$$\sigma^x(y) = \begin{cases} \sigma(y) & \text{if } y \neq x, \\ 1 - \sigma(x) & \text{if } y = x, \end{cases}$$

and

$$\lim_{\beta \to +\infty} -\frac{1}{\beta} \log c_\beta(x,\sigma) = \begin{cases} 1 & \text{if } \sigma(x) = 0 \text{ and } \displaystyle\sum_{y \in \mathscr{N}(x)} \sigma(y) \leqslant 2, \\ \alpha & \text{if } \sigma(x) = 1 \text{ and } \displaystyle\sum_{y \in \mathscr{N}(x)} \sigma(y) \geqslant 2, \\ 0 & \text{otherwise,} \end{cases}$$

for some real parameter $\alpha > 1$ expressing the strength of the bias in favour of the vote for candidate number 1.

7.1. Fast reduction algorithm for the biased majority vote process

Let $\tilde{A}_1 = \{\sigma \in E; H(\{\sigma\}) > 0\} = A_{k_1}$ where $k_1 = |\tilde{A}_1|$. The subdomain $\tilde{A}_1$ is made of the configurations where no voter has more than two neighbours in disagreement with him:

$$\tilde{A}_1 = \left\{ \sigma \in E; \max_{x \in T} \sum_{y \in \mathcal{N}(x)} |\sigma(x) - \sigma(y)| \leqslant 2 \right\}.$$

Lemma 7.1. *The set $\tilde{A}_1$ satisfies*

$$\mathcal{M}(E \setminus \tilde{A}_1) = \{\{\sigma\}; \sigma \in E \setminus \tilde{A}_1\},$$

and therefore is a valid choice for the fast reduction algorithm.

Proof. Let $\sigma \in E \setminus \tilde{A}_1$. We have to build a path $(\sigma_i)_{i=0}^r$ such that $\sigma_0 = \sigma$, $\sigma_r \in \tilde{A}_1$ and $V(\sigma_{i-1}, \sigma_i) = 0$, $i = 1, \ldots, r$. For such a path we have necessarily $\sigma_i = \sigma_{i-1}^{x_i}$ for some $x_i \in T$. Let us build the sequence $(x_i)_{i=1}^r$ in the following way: choose x_i such that

$$(1 - \sigma_{i-1}(x_i)) \sum_{y \in \mathcal{N}(x_i)} \sigma_{i-1}(y) \geqslant 3$$

as long as this is possible. This creates a strictly increasing, and therefore finite sequence $(\sigma_i)_{i=0}^{\ell}$ such that

$$\max_{x \in T} (1 - \sigma_\ell(x)) \sum_{y \in \mathcal{N}(x)} \sigma_\ell(y) \leqslant 2.$$

Now choose $(x_i)_{i=\ell+1}^r$ such that

$$\sigma_{i-1}(x_i) \sum_{y \in \mathcal{N}(x_i)} (1 - \sigma_{i-1}(y)) \geqslant 3$$

this creates a strictly decreasing and therefore finite sequence $(\sigma_i)_{i=\ell}^r$ such that

$$\max_{x \in T} \sigma_r(x) \sum_{y \in \mathcal{N}(x)} (1 - \sigma_r(y)) \leqslant 2.$$

Moreover

$$(1 - \sigma_i(x_i)) \sum_{y \in \mathcal{N}(x_i)} \sigma_i(y) \leqslant 1, \quad i = \ell + 1, \ldots, r,$$

and this shows by induction starting from σ_ℓ that

$$\max_{x \in T} (1 - \sigma_r(x)) \sum_{y \in \mathcal{N}(x)} \sigma_r(y) \leqslant 2.$$

Thus $\sigma_r \in \tilde{A}_1$. $\square$

Lemma 7.2. *Let us put*

$$\tilde{A}_2 = \tilde{A}_1 \setminus \left\{ \sigma \in \tilde{A}_1; \max_{x\in T} (1-\sigma(x)) \sum_{y\in\mathcal{N}(x)} \sigma(y) = 2 \right\},$$

$$\mathcal{R} = \tilde{A}_2 \setminus \{\underline{0}, \underline{1}\}.$$

The configurations in $\tilde{A}_2$ are those where no voter for candidate 0 *has more than one neighbour voting for candidate* 1. *The configurations in $\mathcal{R}$ are made of unions of rectangles of type $\mathbf{1}(\{0,\dots,m-1\}\times\{0,\dots,n-1\})$, with $2\leqslant m\leqslant N-2$ and $2\leqslant n\leqslant N-2$, of stripes of type $\mathbf{1}(\{0,\dots,m-1\}\times\mathbb{Z}/N\mathbb{Z})$ or $\mathbf{1}(\mathbb{Z}/N\mathbb{Z},\times\{0,\dots,n-1\})$ and of loops of type $\mathbf{1}(\{0\}\times\mathbb{Z}/N\mathbb{Z})$ or $\mathbf{1}(\mathbb{Z}/N\mathbb{Z}\times\{0\})$, the mutual distance between any two of these components being not less than* 3.

The heights of the one point cycles of $\mathscr{C}(\tilde{A}_1)$ are the following:

$$H^{\tilde{A}_1}(\{\sigma\}) = \begin{cases} 4\alpha & \text{if } \sigma = \underline{1}, \\ 4 & \text{if } \sigma = \underline{0}, \\ \alpha \wedge 2 & \text{if } \sigma \in \mathcal{R}, \\ 1 & \text{if } \sigma \in \tilde{A}_1 \setminus \tilde{A}_2. \end{cases}$$

Proof. This is an easy consequence of the formula

$$H^{\tilde{A}_1}(\{\sigma\}) = \min\left\{ \sum_{i=1}^{r} V(\sigma_{i-1},\sigma_i); \right.$$
$$\left. \sigma_0=\sigma,\ (\sigma_i)_{i=1}^{r-1} \in (E\setminus\tilde{A}_1)^{r-1},\ \sigma_r\in\tilde{A}_1,\ 1\leqslant r<+\infty \right\}. \qquad \square$$

Lemma 7.3. *A second valid step in the fast reduction algorithm is to choose $A_{|\tilde{A}_2|}=\tilde{A}_2$. Indeed*

$$\mathcal{M}^{\tilde{A}_1}(\tilde{A}_1\setminus\tilde{A}_2) = \{\{\sigma\}; \sigma\in\tilde{A}_1\setminus\tilde{A}_2\}.$$

Proof. Starting from any given $\sigma\in\tilde{A}_1\setminus\tilde{A}_2$, we have to build $(\sigma_i)_{i=0}^r$ such that $\sigma_0=\sigma$, $(\sigma_i)_{i=1}^{r-1}\in(\tilde{A}_1\setminus\tilde{A}_2)$, $\sigma_r\in\tilde{A}_2$ and $V^{\tilde{A}_1}(\sigma_{i-1},\sigma_i)-1=0$. Expressing $V^{\tilde{A}_1}$ in terms of V, we can as well find $(\sigma_i)_{i=0}^r$ such that $\sigma_0=\sigma$, $(\sigma_i)_{i=1}^{r-1}\in(E\setminus\tilde{A}_2)^{r-1}$, $\sigma_r\in\tilde{A}_2$ and $V(\sigma_{i-1},\sigma_i)-\mathbf{1}(\sigma_{i-1}\in(\tilde{A}_1\setminus\tilde{A}_2))=0$. We can do this by putting $\sigma_i=\sigma_{i-1}^{x_i}$, where x_i is chosen such that

$$(1-\sigma_{i-1}(x_i))\sum_{y\in\mathcal{N}(x_i)}\sigma_{i-1}(y) = \max_{x\in T}(1-\sigma_{i-1}(x))\sum_{y\in\mathcal{N}(x)}\sigma_{i-1}(y)\geqslant 2$$

as long as this is possible. This will create a strictly increasing and therefore finite non-self-intersecting sequence of null cost such that $\sigma_r\in\tilde{A}_2$. $\square$

We will give in the next lemma some lower bounds for the large deviation rates of some transitions of the process reduced to $\tilde{A}_2$. Special cases of these transitions include, respectively, sticking a line of ones on the border of a rectangle of ones, removing a line of ones on the border of a rectangle of ones and cutting a stripe of breadth l.

The third estimate is presumably not sharp and will not be used in the following, we give it here because it indicates where we were stopped in our endeavours to compute $V^{A_2}(g_1,g_2)=V^{A_2}(\underline{1},\underline{0})$, the rate of the inverse of the metastable transition.

Lemma 7.4. *In all the following l is any integer in the range $2\leqslant l\leqslant N-2$.*

1. *Let $\sigma,\tilde{\sigma}\in\tilde{A}_2$ be such that*
$$\{0,1\}\times\{0,\ldots,l-1\}\subset\sigma^{-1}(0)$$
and
$$\{1\}\times\{0,\ldots,l-1\}\subset\tilde{\sigma}^{-1}(1),$$
then
$$V^{\tilde{A}_2}(\sigma,\tilde{\sigma})\geqslant 2.$$
2. *Let $\sigma,\tilde{\sigma}\in\tilde{A}_2$ be such that*
$$\{0,1\}\times\{0,\ldots,l-1\}\subset\sigma^{-1}(1),$$
and
$$\{1\}\times\{0,\ldots,l-1\}\subset\tilde{\sigma}^{-1}(0),$$
then
$$V^{\tilde{A}_2}(\sigma,\tilde{\sigma})\geqslant 1+\left\lfloor\frac{l}{2}\right\rfloor(\alpha-1).$$
3. *Let $\sigma,\tilde{\sigma}\in\tilde{A}_2$ be such that*
$$\mathbb{Z}/N\mathbb{Z}\times\{0,\ldots,l-1\}\subset\sigma^{-1}(1),$$
and
$$\{0,1\}\times\{0,\ldots,l-1\}\subset\tilde{\sigma}^{-1}(0),$$
then
$$V^{\tilde{A}_2}(\sigma,\tilde{\sigma})\geqslant 2\alpha.$$

Proof. We will use the formula

$$V^{\tilde{A}_2}(\sigma,\tilde{\sigma})=\min\Bigg\{V(\sigma_0,\sigma_1)+\sum_{i=2}^{r}V(\sigma_{i-1},\sigma_i)-\mathbf{1}(\sigma_{i-1}\in\tilde{A}_1);$$
$$\sigma_0=\sigma,\ \sigma_r=\tilde{\sigma},\ (\sigma_i)_{i=1}^{r-1}\in(E\setminus\tilde{A}_2)^{r-1},\ \sigma_i=\sigma_{i-1}^{x_i},\ 1\leqslant r<+\infty\Bigg\}.$$

1. As $\sigma_0\in\tilde{A}_2$, then $V(\sigma_0,\sigma_1)\geqslant 1$. Let
$$j=\inf\{i;|\sigma_i^{-1}(1)\cap\sigma^{-1}(0)|=2\},$$
then $\sigma_{j-1}\in E\setminus\tilde{A}_1$ (because $\sigma_{j-1}^{-1}(1)\cap\sigma^{-1}(0)=\{x\}$ such that $\sum_{y\in\mathcal{N}(x)}\sigma(y)\leqslant 1$), and $V(\sigma_{j-1},\sigma_j)=1$, thus $V^{\tilde{A}_2}(\sigma,\tilde{\sigma})\geqslant 2$.

2. Let $j = \inf\{i; x_i \in \sigma^{-1}(1)\}$, then $\sigma_{j-1}(x_j) = 1$ and $\sum_{y \in \mathcal{N}(x_j)} \sigma_{j-1}(y) \geqslant 2$, therefore $V(\sigma_{j-1}, \sigma_j) - \mathbf{1}(\sigma_{j-1} \in \tilde{A}_1) \geqslant \alpha - 1$. Moreover if $j > 1$, then $V(\sigma_0, \sigma_1) = 1$, and if $j = 1$ then $V(\sigma_0, \sigma_1) = \alpha$, thus

$$V(\sigma_0, \sigma_1) + \sum_{i=2}^{j} V(\sigma_{i-1}, \sigma_i) - \mathbf{1}(\sigma_{i-1} \in \tilde{A}_1) \geqslant \alpha.$$

Now let us consider the "stable" structures

$$Q_s = \{0, 1\} \times \{2s, 2s+1\}, \quad s = 0, \dots, \left\lfloor \frac{l}{2} \right\rfloor - 1,$$

and let

$$j_s = \inf\{i; x_i \in Q_s\}.$$

We have $j_s \leqslant r$ because $Q_s \subset \sigma^{-1}(1)$ and $Q_s \cap \tilde{\sigma}^{-1}(0) \neq \emptyset$. Moreover the sets Q_s being disjoint, the j_s are distinct indices. We have from the definition of j_s that $Q_s \subset \sigma_{j_s-1}^{-1}(1)$ and therefore that $V(\sigma_{j_s-1}, \sigma_{j_s}) - \mathbf{1}(\sigma_{j_s-1} \in \tilde{A}_1) \geqslant \alpha - 1$.
The other terms in the sum being nonnegative, we have

$$\begin{aligned} &V(\sigma_0, \sigma_1) + \sum_{i=2}^{r} V(\sigma_{i-1}, \sigma_i) - \mathbf{1}(\sigma_{i-1} \in \tilde{A}_1) \\ &\geqslant \begin{cases} \alpha + \lfloor \frac{l}{2} \rfloor (\alpha - 1) & \text{if } j < \inf\{j_s\}, \\ 1 + \lfloor \frac{l}{2} \rfloor (\alpha - 1) & \text{if } j = \inf\{j_s\}, \end{cases} \\ &\geqslant 1 + \left\lfloor \frac{l}{2} \right\rfloor (\alpha - 1). \end{aligned}$$

3. Let

$$j_1 = \inf\{i; x_i \in \mathbb{Z}/N\mathbb{Z} \times \{0, \dots, l-1\}\},$$

then

$$\begin{aligned} V(\sigma_0, \sigma_1) + \sum_{i=2}^{j_1} V(\sigma_{i-1}, \sigma_i) - \tilde{H}(\{\sigma_{i-1}\}) &\geqslant \begin{cases} 1 + (\alpha - 1) & \text{if } j_1 > 1, \\ \alpha & \text{if } j_1 = 1, \end{cases} \\ &\geqslant \alpha. \end{aligned}$$

Let

$$j_2 = \inf\{i; |\sigma_i^{-1}(0) \cap \mathbb{Z}/N\mathbb{Z} \times \{0, \dots, l-1\}| = 2\}$$

then $\sigma_{j_2-1} \notin \tilde{A}_1$ and $j_2 > j_1$, therefore

$$\sum_{i=j_1+1}^{j_2} V(\sigma_{i-1}, \sigma_i) - \tilde{H}(\{\sigma_{i-1}\}) \geqslant V(\sigma_{j_2-1}, \sigma_{j_2}) \geqslant \alpha,$$

and $V^{\tilde{A}_2}(\sigma, \tilde{\sigma}) \geqslant 2\alpha$. $\square$

Lemma 7.5. *Let*

$$\ell(\{0,\dots,m-1\}\times\{0,\dots,n-1\}) = \min\{m,n\},\quad 2\leqslant m\leqslant N-2,\ 2\leqslant n\leqslant N-2,$$

$$\ell(\{0,\dots,m-1\}\times\mathbb{Z}/N\mathbb{Z}) = +\infty,\quad 2\leqslant m\leqslant N-2,$$

$$\ell(\{0\}\times\mathbb{Z}/N\mathbb{Z}) = 2,$$

let the function ℓ be invariant with respect to translations and the symmetry with respect to the first diagonal, let us extend ℓ to $\mathscr{R}$ in the unique way that satisfies

$$\ell(\sigma\vee\sigma') = \min\{\ell(\sigma),\ell(\sigma')\},\quad \sigma,\sigma',\sigma\vee\sigma'\in\mathscr{R}.$$

Then we have

$$H^{\tilde{A}_2}(\{\sigma\}) = \min\left(2,\left\lfloor\frac{\ell(\sigma)}{2}\right\rfloor(\alpha-1)+1\right),\quad \sigma\in\mathscr{R},$$

$$H^{\tilde{A}_2}(\{\underline{0}\}) = 4,$$

$$H^{\tilde{A}_2}(\{\underline{1}\}) = 4\alpha.$$

Proof. We have from the previous lemma that

$$V^{\tilde{A}_2}(\mathbf{1}(\{0,\dots,m-1\}\times\{0,\dots,n-1\}),\mathbf{1}(\{0,\dots,m\}\times\{0,\dots,n-1\}))\geqslant 2,$$

considering an explicit (and obvious) suitable path, we see that equality holds in this formula. In the same way

$$\begin{aligned}&V^{\tilde{A}_2}(\mathbf{1}(\{0,\dots,m-1\}\times\{0,\dots,l-1\}),\mathbf{1}(\{0,\dots,m-2\}\times\{0,\dots,l-1\}))\\&\quad=(\alpha-1)\left\lfloor\frac{l}{2}\right\rfloor+1,\end{aligned}$$

(consider the path associated with the sequence of sites $x_i = (m-1,2i-1)$, $i = 1,\dots,\lfloor l/2\rfloor$) and

$$V^{\tilde{A}_2}(\mathbf{1}(\mathbb{Z}/N\mathbb{Z}\times\{0,\dots,l-1\}),\mathbf{1}(\mathbb{Z}/N\mathbb{Z}\times\{0,\dots,l\})) = 2.$$

(The minimising path in this last equation is the obvious one.) □

Theorem 7.1. *The set $\tilde{A}_3=\{\underline{0},\underline{1}\}=A_2$ is a valid choice in the fast reduction algorithm*:

$$\mathscr{M}(\tilde{A}_2\setminus\tilde{A}_3) = \{\{\sigma\};\sigma\in\tilde{A}_2\setminus\tilde{A}_3\}.$$

The state $\underline{0}$ is the unique metastable state of the energy landscape, which therefore behaves as a two well potential.

Moreover the configurations in $\sigma\in\tilde{A}_2\setminus\{\underline{0},\underline{1}\}$ are subcritical if each rectangle $\chi\leqslant\sigma$ satisfies $\lfloor\ell(\chi)/2\rfloor<1$ and supercritical if at least one rectangle $\chi\leqslant\sigma$ satisfies $\lfloor\ell(\chi)/2\rfloor>1$. (*By subcritical we mean that*

$$\lim_{\beta\to+\infty}\mathbb{P}_\sigma^\beta(X_{\tau(E\setminus\{\underline{0},\underline{1}\})}=\underline{0}) = 1,$$

and by supercritical we mean that

$$\lim_{\beta\to+\infty}\mathbb{P}_\sigma^\beta(X_{\tau(E\setminus\{\underline{0},\underline{1}\})}=\underline{1}) = 1.)$$

Proof. Any configuration in $\sigma \in \tilde{A}_2 \setminus \tilde{A}_3$ can be shrunk to $\underline{0}$ (if each of its rectangles χ is such that $\lfloor \ell(\chi)/2 \rfloor(\alpha - 1) \leqslant 1$) or to $\underline{1}$ (if one of its rectangles χ is such that $\lfloor \ell(\chi)/2 \rfloor(\alpha - 1) \geqslant 1$) through a path $(\sigma_i)_{i=0}^{r}$ such that $(\sigma_i)_{i=1}^{r-1} \in (\tilde{A}_2 \setminus \tilde{A}_3)^{r-1}$ and

$$V^{\tilde{A}_2}(\sigma_{i-1}, \sigma_i) - H^{\tilde{A}_2}(\{\sigma_{i-1}\}) = 0, \quad i = 1, \ldots, r.$$

This can easily be seen from the previous computations. That $\underline{0}$ is the unique metastable state comes from the fact that

$$H^{\tilde{A}_2}(\underline{0}) > \sup_{\sigma \in \tilde{A}_2 \setminus \{\underline{0}, \underline{1}\}} H^{\tilde{A}_2}(\sigma). \qquad \square$$

7.2. Some facts about the exit path

The next natural step would be to compute $V^{A_2}(\underline{0}, \underline{1})$ and $V^{A_2}(\underline{1}, \underline{0})$. This would involve solving the minimisation problems

$$\min \left\{ V(\sigma_0, \sigma_1) + \sum_{i=2}^{r} V(\sigma_{i-1}, \sigma_i) - \tilde{H}(\sigma_{i-1}); \right.$$
$$\left. \times \sigma_0 = \eta, \sigma_r = \zeta, (\sigma_i)_{i=0}^{r} \in \mathscr{V}_2^{|E|} \right\}, \quad (\eta, \zeta) \in \{(\underline{0}, \underline{1}), (\underline{1}, \underline{0})\}.$$

Unfortunately these problems seem to be difficult to tackle.

To give an idea of the difficulties involved, let us concentrate on the computation of $V^{A_2}(\underline{0}, \underline{1})$. Any corresponding minimising path $(\hat{\sigma}_i)_{i=0}^{\hat{r}}$ gives a description of one typical "loop erased" exit path (one with zero cost or equivalently nonvanishing probability to occur at low temperature).

Let (by analogy with the Ising model, see Neves and Schonmann, 1991, 1992; Schonmann, 1992; Alonzo and Cerf, 1996; Ben Arous and Cerf, 1996) I be the set of reduction paths with the following structure (up to translations and other obvious cost invariant transformations):

$$\begin{aligned}
&\underline{0},\\
&\mathbf{1}(\{1,2\}^2),\\
&\mathbf{1}(\{1,2\} \times \{1,2,3\}),\\
&\mathbf{1}(\{1,2,3\}^2),\\
&\mathbf{1}(\{1,2,3\} \times \{1,2,3,4\}),\\
&\ldots\\
&\underline{1}.
\end{aligned} \tag{7}$$

This is a first naïve guess about what $\hat{\sigma}^{\tilde{A}_2}$ could be. The cost of such paths is

$$4 + \sum_{l=1}^{+\infty} 4(1 - l(\alpha - 1)) \vee 0 = 4 + 2 \left\lfloor \frac{1}{\alpha - 1} \right\rfloor \left(3 - \alpha - \left\lfloor \frac{1}{\alpha - 1} \right\rfloor (\alpha - 1) \right).$$

Unfortunately, we can see that the more direct path

$$
\begin{aligned}
&\underline{0},\\
&\mathbf{1}(\{1,2,3\}^2),\\
&\mathbf{1}(\{1,2,3\}\times\{1,2,3,4\}),\\
&\dots\\
&\underline{1},
\end{aligned}
$$

where the first step is done through $\mathbf{1}(\{(i,j);1\leqslant i\leqslant 3,1\leqslant j\leqslant 3,i+j\in 2\mathbb{N}\})\in E\setminus\tilde{A}_1$, is more probable when $1<\alpha<1.5$, since its cost is

$$
\begin{aligned}
&5+2(2-\alpha)+\sum_{l=2}^{+\infty}4(1-l(\alpha-1))\vee 0\\
&\quad=(2\alpha-3)+4+2\left\lfloor\frac{1}{\alpha-1}\right\rfloor\left(3-\alpha-\left\lfloor\frac{1}{\alpha-1}\right\rfloor(\alpha-1)\right).
\end{aligned}
$$

This is not still a minimal path when α is close to one, indeed the path going through the following states has a lower cost:

$$
\begin{aligned}
&\underline{0},\\
&\mathbf{1}(\{(i,j);1\leqslant i\leqslant 3,1\leqslant j\leqslant 3,i+j\in 2\mathbb{N}\}),\\
&\mathbf{1}(\{1,2,3\}^2),\\
&\mathbf{1}(\{1,2,3\}^2)\vee\mathbf{1}(\{(i,j);4\leqslant i\leqslant 6,4\leqslant j\leqslant 6,i+j\in 2\mathbb{N}\}),\\
&\mathbf{1}(\{1,2,3\}^2)\vee\mathbf{1}(\{4,5,6\}^2),\\
&\mathbf{1}(\{1,\dots,6\}^2),\\
&\dots\\
&\underline{1},
\end{aligned}
$$

its cost being

$$
10-\alpha+\sum_{l=3}^{+\infty}4(1-l(\alpha-1))\vee 0,
$$

which is lower than the previous cost when $1<\alpha<11/9$.

More generally it seems that when α is close to one it is more efficient to merge blocks. We can see for instance that it costs less to merge a 3×3 block at one angle of a square than to add 3 lines and 3 columns in an Ising like way as long as the size of the square is less than approximately $2/5$ of the critical length $2/(\alpha-1)$. Anyhow, there is a variety of possible ways to merge blocks and we were not able to make an exhaustive list of them or to prove that the above one is indeed optimal. Therefore we had to be content with the following theorem (which could still be a little sharpened near $\alpha=1$ from the previous remark):

Theorem 7.2. *The relaxation time from the unique metastable state $\underline{0}$ satisfies the following inequalities*:

$$\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{E}^{\beta}_{\underline{0}}(\tau(E \setminus \{\underline{1}\})) \leqslant 4 + 2 \left\lfloor \frac{1}{\alpha-1} \right\rfloor \left(3 - \alpha - \left\lfloor \frac{1}{\alpha-1} \right\rfloor (\alpha-1)\right) \vee 0$$
$$- \begin{cases} 14 - 11\alpha & \text{if } 1 < \alpha \leqslant 11/9, \\ 3 - 2\alpha & \text{if } 11/9 \leqslant \alpha \leqslant 3/2, \\ 0 & \text{if } 3/2 \leqslant \alpha. \end{cases}$$

Remark 7.1. We can also use the large deviation principle stated in Theorem 6.1 to prove that the exit path is different from the typical exit paths I of the Ising model at low temperature when the bias parameter α is between 1 and 1.5. Namely, with the notation introduced in Eq. (7), we have that

$$\lim_{\beta\to+\infty} \frac{1}{\beta} \log \mathbb{P}^{\beta}_{\underline{0}}(\Gamma((X_i)_{i=\theta}^{\tau}) \in I) \leqslant - \begin{cases} 14 - 11\alpha & \text{if } 1 < \alpha \leqslant 11/9, \\ 3 - 2\alpha & \text{if } 11/9 \leqslant \alpha \leqslant 3/2. \end{cases}$$

This result is qualitatively interesting, it proves that the biased majority vote process is more likely to build first checker board patterns of votes for candidate number one, than to build steadily increasing packed square patterns when the bias parameter α is not too strong.

8. Conclusion

We have shown on the example of the biased majority vote process on the torus that the reduction method gives a systematic way to find the metastable state and to compute precise informations about the "loop erased" exit path of the chain. Moreover, the reduced dynamics and their transition rates give a picture of what is going on at different exponential time scales.

Appendix A. Some comments on related approaches

Other algorithms to compute the critical depth and analyse the exit path include one proposed by Trouvé (1994) and one proposed by Olivieri and Scoppola (1996).

Trouvé's algorithm computes recursively cycles of increasing depths, and is therefore quite different from the reduction method, that allows to get the free energy first with the possibility to deduce the full cycle decomposition from it afterwards if needed (from Eq. (4)), but also the possibility to compute only the deepest cycles in cases when the full decomposition would be intractable. This is what happened to us with the vote process: in principle, we could have followed Trouvé's algorithm, but this proved to be beyond our computing skill. We designed the reduction method to make computations simpler.

The renormalisation procedure of Olivieri and Scoppola builds a succession of Markov chains on quotient state spaces. They remove at each step unstable states, group together strongly connected stable equilibria, and renormalise time to create new grouping or removal opportunities at the next step. To each state at one level corresponds a set of states at the previous level. The chains are defined by their transition rates on different probability spaces. Moreover Olivieri and Scoppola restrict their study to most probable events. To describe the most probable exit paths, they trace back everything to the original Markov chain using a notion of "path refinement" that leads to some special kind of exit path "tubes". They do not compute the virtual energy.

There is some analogy however between the renormalisation and the reduction approaches: their aim is the same, to study sets (or "tubes") of trajectories solving a succession of path minimisation problems. The reduction method is a simpler construction, made on a single probability space based on a single state space, using hitting times only, and not involving computations of strongly connected components nor any time renormalisation (so that all our chains have the same critical depth). It makes it easier to push the study of the exit path further than the characterisation of most probable events, leads to the simple and natural notion of loop erased exit paths, and produces large deviation estimates for exit path events with exponentially small probabilities. This is used in the application, where we were not able to compute explicitely the most probable loop erased exit path, but were still able to show that the most obvious candidate (namely the most probable loop erased exit path for the Ising model) has in fact an exponentially small probability to occur.

One source of inspiration for the reduction method was the structure of random times used in the proof of speed of convergence estimates for nonhomogeneous Markov chains in Catoni (1992). In this paper the attractors are ordered according to their depths, and working backward in time, the last visit to the deepest secondary attractor is introduced, then the last visit to the deepest remaining one, etc., to grasp the influence of the cooling schedule on smaller and smaller time scales.

9. For further reading

The following references are also of interest to the reader: Chen et al., 1995; Durrett, 1988; Deuschel and Mazza 1994; Scoppola, 1993; Trouvé, 1993; Trouvé, 1995

References

Alonzo, L., Cerf, R., 1996. The three dimensional polyominoes of minimal area. Electron. J. Combin. 3(1).

Ben Arous, G., Cerf, R., 1996. Metastability of the three dimensional Ising model on a Torus at very low temperatures. Electron. J. Probab. 1(10), 55.

Catoni, O., 1992. Rough large deviation estimates for simulated annealing: application to exponential schedules. Ann. Probab. 20(3), 1109–1146.

Catoni, O., 1998. Simulated annealing algorithms and Markov chains with rare transitions. Sém. Probab. to appear, 1–51, in French (1995), English preprint at http://www.dmi.ens.fr/preprints (1997), revision at http://www.proba.jussieu.fr/users/catoni/lecture5.dvi (1998).

Catoni, O., Cerf, R., 1997. The exit path of a Markov chain with rare transitions. ESAIM: P & S, Vol. 1, pp. 95–144, http://www.emath.fr/ps/

Chen, D., 1997. The consensus times of the majority vote process on a torus. J. Statist. Phys. 86(3/4), 779–802.

Chen, D., Feng, J., Qian, M.P., 1995. The metastable behavior of the two-dimensional Ising model. In: Ma, Z., Röckner, M., Yan, J. (Eds.), The Proceedings of the International Conference on Dirichlet Forms and Stochastic Processes. Walter de Gruyter, Berlin, pp. 73–86.

Chen, D., Feng, J., Qian, M.P., 1996. The metastability of exponentially perturbed Markov chains. Sci. China, Ser. A, English Edition 39, 7–28.

Deuschel, J-D., Mazza, C., 1994. L^2 convergence of time nonhomogeneous Markov process: I. spectral estimates. Ann. Appl. Probab. 4, 1012–1056.

Durrett, R., 1988. Lecture Notes on Particle Systems and Percolation. Wadsworth & Brooks/Cole, Monterey, California.

Freidlin, M.I., Wentzell, A.D., 1984. Random Perturbations of Dynamical Systems. Springer, New York.

Hajek, B., 1988. Cooling schedule for optimal annealing. Math. Oper. Res. 13, 311–329.

Kirchhoff, G., 1847. Über die Auflösung der Gleichungen, auf welche man beider Untersuchung der linearen Verteilung galvanischer Ströme gefuhrt wird. Ann. Phys. Chem. 72, 497–508. (English transl. IRE Trans. Circuit Theory CT-5 (1958) 4–7).

Neves, E.J., Schonmann, R.H., 1991. Critical droplets and metastability for a Glauber dynamics at very low temperatures. Comm. Math. Phys. 137(2), 209–230.

Neves, E.J., Schonmann, R.H., 1992. Behavior of droplets for a class of Glauber dynamics at very low temperature. Probab. Theory Related Fields 91(3–4), 331–354.

Olivieri, E., Scoppola, E., 1996. Markov chains with exponentially small transition probabilies: first exit problem from a general domain II. The general case. J. Statist. Phys. 84(5/6), 987–1041.

Schonmann, R.H., 1992. The pattern of escape from metastability of a stochastic Ising model. Commun. Math. Phys. 147, 231–240.

Scoppola, E., 1993. Renormalization group for Markov chains and application to metastability. J. Statist. Phys. 73, 83–122.

Trouvé, A., 1993. Parallélisation massive du recuit simulé. Doctoral Dissertation, Université Paris 11, January 5.

Trouvé, A., 1994. Cycle decomposition and simulated annealing. SIAM J. Control Optim. 34(3), 1996.

Trouvé, A., 1995. Rough large deviation estimates for the optimal convergence speed exponent of generalized simulated annealing algorithms. Ann. Inst. H. Poincaré, Probab. Statist. 32(2), 1996.